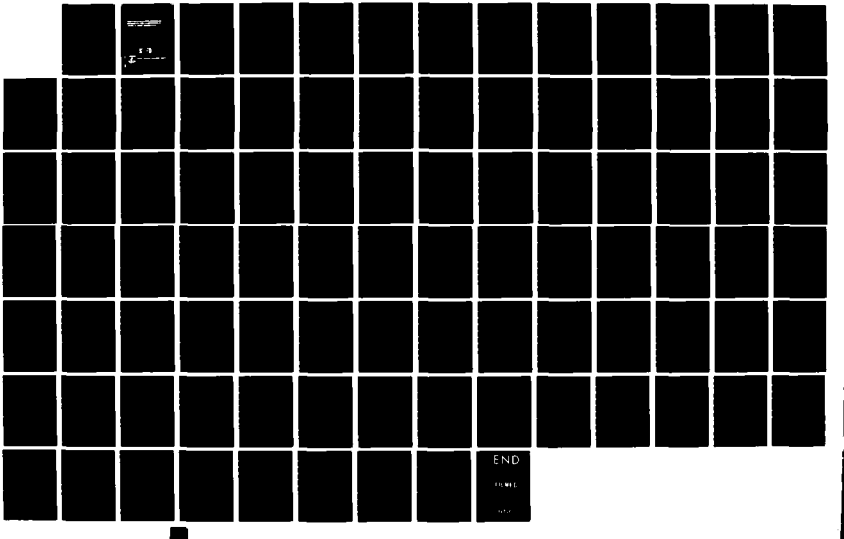
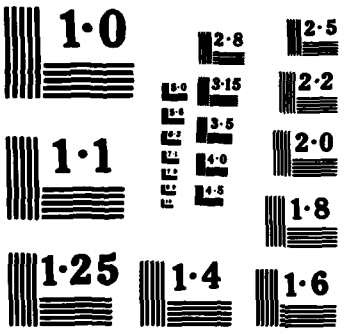
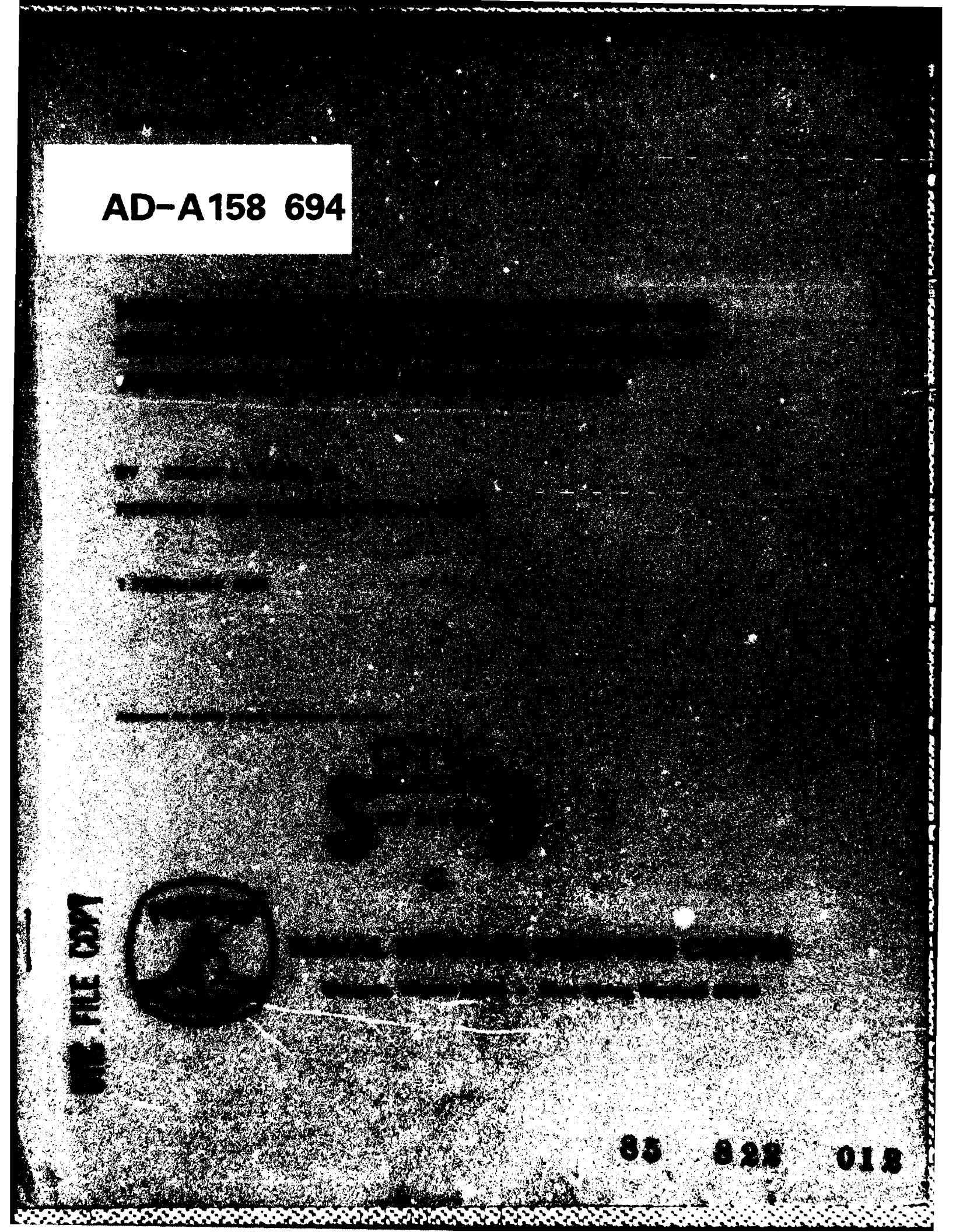
AD-A158 694 HIERARCHICAL RELAXATION METHODS FOR MULTISPECTRAL PIXEL 1/1
CLASSIFICATION AS..(U) NAVAL SURFACE WEAPONS CENTER
SILVER SPRING MD E A COHEN 01 FEB 85 NSWC/TR-84-54
UNCLASSIFIED F/G 17/8 NL





NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: NSWC TR 84-54	2. GOVT ACCESSION NO.: AD-A158 694
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)*: HIERARCHICAL RELAXATION METHODS FOR MULTISPECTRAL PIXEL CLASSIFICATION AS APPLIED TO TARGET IDENTIFICATION	5. TYPE OF REPORT & PERIOD COVERED: Final
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Edgar A. Cohen, Jr.	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Surface Weapons Center (Code R44), 10901 New Hampshire Avenue, White Oak, Silver Spring, MD 20903-5000	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE: 1 February 1985
	13. NUMBER OF PAGES: 91
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*	15. SECURITY CLASS. *(of this report)*: UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Image Modeling
Wave Bands;
Split and merge algorithm;
Generalized sloped facet model;
Sobolev model;
Kalman filter;
Nonlinear filtering;
Boundary classification;
Moment area method;
Relaxation methodology;
Stochastic labeling;
Quad tree;
Interesting components;
Background replacement concept.

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

This report provides ~~some new~~ insights into the approaches toward image modeling as applied to target detection, whether that target be a ship, an incoming missile, or an aircraft. The approach ~~taken~~ is that of examining the energy in prescribed wave-bands which emanate from a target and correlating the emissions. Typically, one might be looking at two or three infrared bands, possibly together with several visual bands, ~~as well~~. The target is segmented, using both first and second order modeling, into a set of "interesting

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

Block # 19 - Key Words (Cont'd)
Multispectral classification

Block # 20 - Abstract (Cont'd)

components," and these components are correlated ~~in such manner~~, as to enhance the classification process.

A Markov-type model is used to provide an a priori assessment of the spatial relationships among critical parts of the target, and a stochastic model using the output of an initial probabilistic labeling is invoked. The tradeoff between this stochastic model and the Markov model is then optimized to yield a best labeling for identification purposes. In an identification of friend or foe (IFF) context, the methodology presented in this report could be of interest, for it provides the ingredients for such a higher level of understanding.

This work was previously supported under Independent Research funds at NSWC.

FOREWORD

This report provides some new insights into and approaches toward image modeling as applied to target identification, whether that target be a ship, an incoming missile, or an aircraft. The approach taken is that of examining the energy in prescribed wave-bands which emanates from a target and correlating the emissions. Typically one might be looking at two or three infrared bands, together possibly with several visual bands as well. The target is segmented, using both first and second order modeling, into a set of "interesting components," and these components are correlated in such manner as to enhance the classification process.

A Markov-type model is used to provide an a priori assessment of the spatial relationships among critical parts of the target, and a stochastic model using the output of an initial probabilistic labeling is invoked. The tradeoff between this stochastic model and the Markov model is then optimized so as to yield a best labeling for identification purposes. In an identification of friend or foe (IFF) context, the methodology presented in this report could be of great interest, for it provides the ingredients for such a higher level of understanding.

We wish to thank several people who have provided both ideas and encouragement for this investigation. In particular, Mr. Gordon Powell and Mr. Kent Anderson, formerly of NSWC, provided interesting comments. Mr. Powell, acquainted the author with much of the literature in the field. Also, Mr. John Moscar and Mr. Mike Rowan were very helpful in the early stages of the effort, since through them the author became familiar with previous applied work at NSWC. With regard to infrared processing, the author has had an interesting working relationship with Mr. Donald M. Wilson and has learned much about the benefits of infrared sensing in target discrimination from him. Finally, the author would like to recognize Dr. John Wingate, Dr. Brooke Stephens, and Dr. K. Y. Chien, of R44, who provided interesting mathematical and physical insights.

This work was previously supported under Independent Research funds at NSWC.

Approved by:

Ira M Blatstein

IRA M. BLATSTEIN, Head
Radiation Division

CONTENTS

ILLUSTRATIONS

CHAPTER 1

INTRODUCTION

In recent work on aircraft identification [1,2] and ship classification, the classical theory of invariants [3] was employed in order to classify a target known to belong to one of a given number of target classes. The classes were constructed in the following manner: A target prototype from the class was viewed from a number of perspectives and under ideal imaging conditions. Then, thinking of the two-dimensional projections of the target as binary images, the moment area method was used to construct a feature vector for the given perspective. Each feature, being a proper combination of moments, had three fundamental invariance properties, which were invariance to position of the image in the field of view, invariance to rotation in the image plane, and invariance to range of the target. Of course, a real target is a superposition of an ideal target and background noise due to environmental effects. In addition, noise will be introduced by the imaging system itself. Therefore, before one can successfully apply any pattern recognition method, he must be able to extract this noisy target from the background. There are two steps to this process: (1) The picture is subjected to some kind of smoothing operation, so that, hopefully, the noise is reduced to a manageable level. (2) Now that one has a sensible picture, one extracts the target by boundary identification, which often requires some kind of gradient algorithm. Once the object has been extracted, it is compared to the data base of ideal target prototypes, and the object is assigned to that class which is closest, in some well-defined sense, to the image representation.

We would like to emphasize at this point that previous work has often been centered around a rather naive interpretation of the object of interest. Each image, be it optical, infrared, or otherwise, is conceived in terms of two color levels, black and white, with black being assigned a 1 and white a zero (or vice versa). Such a two-level description is deemed appropriate from the point of view of computational feasibility, provided that all that is really necessary anyway is a crude target assignment. However, in practice, it is easy to imagine a situation in which a higher level description is not only desirable, but indeed necessary. For example, it is one thing to determine that a given ship is a destroyer, rather than an aircraft carrier or a tanker, but it may be

quite another to ascertain whether or not it is an enemy or a friendly destroyer. (Imagine a combat situation in which such discrimination is crucial.)

The use of relaxation methods for target discrimination seems to be ideally suited to the task at hand. In particular, the work of Rosenfeld, Hummel and Zucker [4] appears to be of fundamental interest. One sees immediately that such methods should engender a higher-level definition of the object if one can appropriately define the components of interest, together with a meaningful assignment of their interrelationships. However, what one does not usually see in the literature is a marriage of the concepts of relaxation and multispectral classification. This would seem to be a natural extension, and one of the purposes of this document will be to investigate the potential of such a union. In addition, we shall pursue the theory of contour or boundary extraction in the context of multispectral information. We note that some work in this direction has been pursued recently by Eklundh, Yamamoto, and Rosenfeld,[5] and that the results have been quite encouraging. Also, we shall invoke the work of Haralick,[6] Nahi,[7] Kaufman,[8] and Woods [9] and attempt to unify their concepts in a meaningful way.

Our analysis will begin with a procedure similar to that of Chen and Pavlidis.[10] That is to say, we shall show how to devise a split-and-merge algorithm for our purpose. Implicit in this methodology is the incorporation of the idea of the quad tree.[11]

It is hoped that this work will prove useful in more carefully defining the structure of targets.

CHAPTER 2

THE GENERALIZED SPLIT-AND-MERGE ALGORITHM

We shall describe here in general terms what we propose to do with regard to a unification of the work of Haralick [6] and Chen and Pavlidis.[10] Haralick's sloped facet model is to be employed in conjunction with the splitting phase of the Chen and Pavlidis procedure. That is one new idea. Secondly, we shall superimpose a color (or pseudocolor) scheme for two reasons: (1) It is well-known that, by so doing, one may make the target intensity profile insensitive to changes in illumination.[12] This is a distinct advantage from a physical point of view. (2) In infrared detection, where our methods may prove most valuable, it is important to examine gray-level intensity in different wave bands and to correlate the results so obtained. Such methodology has proved to be of great value in LANDSAT image studies.

Now let us begin with the idea of the quad tree. Suppose that we have a rectangular target area, as shown in Figure 2-1. We shall divide this rectangle into four congruent subrectangles as shown. Now let us assume that our target pixels are described in terms of normalized primary colors (or their equivalents). That is, we have the nonnegative vector $(g_R, g_B, g_G)^t$, with $g_R + g_B + g_G = 1$. Then let us compute the sample covariance matrix [13] for the entire rectangle and compare it with those for the subrectangles. Suppose that K is the overall covariance matrix and that K_i, $1 \leq i \leq 4$, is the matrix for subrectangle i, as indicated in the figure. The game will now be

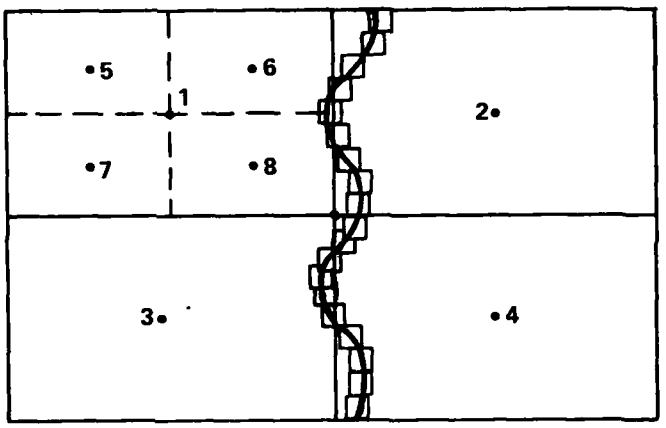


FIGURE 2-1. THE QUAD-TREE SPLIT-AND-MERGE ALGORITHM

played as follows: If the maximum norm of the difference between K and K_i, i.e., $\max_{1 \leq i \leq 4} ||K - K_i||$, does not exceed some small positive , then we proceed with Haralick's method for determining whether or not all the subrectangles belong to the same sloped surface (facet). On the other hand, when the max of the norms does not pass the epsilon test, we do not use Haralick's procedure. In that case, we subdivide rectangles 1, 2, 3, and 4 into subrectangles. Eventually we shall arrive at a dissection for which the equal covariance assumption is valid. From that point onward the scheme is applicable. Boundary determination should fall out of this process naturally in terms of small boxes containing the boundary points, as illustrated in Figure 2-1.

What we really have in mind is the determination of a number of pertinent components of a target which are useable in later identifying it. The idea is that, once we know the boundaries, we automatically have the regions, which are formed by merging of subregions having the same parameters associated with them. This is the "merge" part of the algorithm. The next chapter is devoted to the modification of Haralick's scheme.

CHAPTER 3
THE GENERALIZED SLOPED FACET MODEL

In what follows we shall make use of the quad-tree concept,[11] whereby a rectangle is divided into four equal parts and each part is again divided into four parts and so on, until a certain group of criteria are met. As shown in Figure 3-1, let us assume that we have a rectangular lattice of points and that the equal covariance assumption of Chapter 2 is valid. Point 0 is the center of our large rectangle, and point 0_i is the center of subrectangle i. Note that all four subrectangles meet at 0. We want to test the hypothesis that all

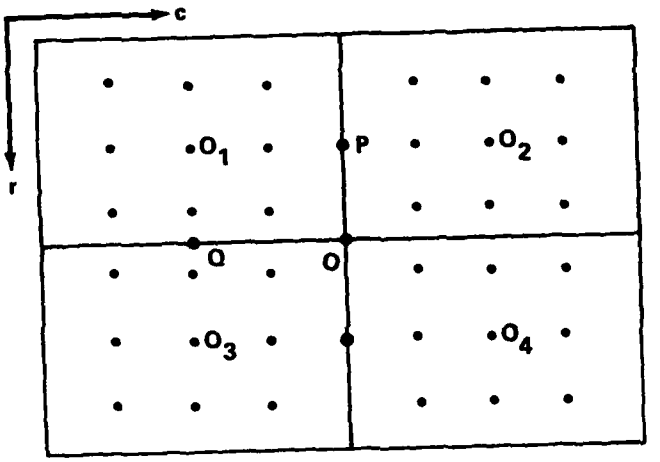


FIGURE 3-1. TESTING THE SLOPED FACET MODEL ASSUMPTIONS

subrectangles are associated with the same sloped planes. We are assuming, therefore, a regression model of the following type:

$$g_R(r,c) = \alpha_R r + \beta_R c + \gamma_R + \eta_R(r,c)$$

$$g_B(r,c) = \alpha_B r + \beta_B c + \gamma_B + \eta_B(r,c) \qquad (3\text{-}1)$$

$$g_G(r,c) = \alpha_G r + \beta_G c + \gamma_G + \eta_G(r,c),$$

where $g_R + g_B + g_G = 1$, r is the row index, and c is the column index. Of course, the constraint that all three color components sum to one allows us to incorporate into the state vector any two of them. We shall assume that,

although the red, blue and green components may be statistically correlated at a given point in the image, they are independent from point to point. Note that this is a first-order assumption, which is given for computational convenience; in fact, a higher order model will be given later in this report. Selecting the red and blue quantities in (3-1), we present the matrix formulation

$$\begin{pmatrix} g_R \\ g_B \end{pmatrix} = \begin{pmatrix} \alpha_R & \beta_R & \gamma_R \\ \alpha_B & \beta_B & \gamma_B \end{pmatrix} \begin{pmatrix} r \\ c \\ 1 \end{pmatrix} + \begin{pmatrix} \eta_R \\ \eta_B \end{pmatrix} \tag{3-2}$$

In the usual way, assuming that K_η is the 2 by 2 covariance matrix for the noise term in (3-2), we may formulate the least squares problem of minimizing

$$\sum_{r \epsilon R} \sum_{c \epsilon C} \hat{\underline{e}}_{r,c}^{\,t} K_\eta^{-1} \hat{\underline{e}}_{r,c}, \tag{3-3}$$

where t denotes transpose of the error vector $\underline{e}_{r,c}$, R is the row index set for the rectangle, and C is the column index set. Furthermore,

$$\hat{\underline{e}}_{r,c} = \begin{pmatrix} g_R \\ g_B \end{pmatrix} - \begin{pmatrix} \hat{\alpha}_R & \hat{\beta}_R & \hat{\gamma}_R \\ \hat{\alpha}_B & \hat{\beta}_B & \hat{\gamma}_B \end{pmatrix} \begin{pmatrix} r \\ c \\ 1 \end{pmatrix}, \tag{3-4}$$

where the matrix in (3-4) is the array of expected coefficients. Note now the fact that K_η^{-1} is symmetric and positive definite, so that it may be diagonalized by an orthogonal matrix P:

$$D_\eta = P K_\eta^{-1} P^t, \tag{3-5}$$

where $D_\eta = \text{diag}(\lambda_1, \lambda_2)$ and λ_1, λ_2 are positive. We are here assuming, of course, that the process is ergodic[14] so that K_η^{-1} is estimable by means of the spatial information given in the gray level profiles.

Substituting (3-5) into (3-3), one has to minimize

$$\sum_{r \epsilon R} \sum_{c \epsilon C} \hat{\underline{f}}_{r,c}^{\,t} D_\eta \hat{\underline{f}}_{r,c}, \tag{3-6}$$

where $\hat{\underline{f}}_{r,c} = P \hat{\underline{e}}_{r,c}$. Thus

$$\hat{\underline{f}}_{r,c} = P\left[\begin{pmatrix} g_R \\ g_B \end{pmatrix} - \begin{pmatrix} \hat{\alpha}_R & \hat{\beta}_R & \hat{\gamma}_R \\ \hat{\alpha}_B & \hat{\beta}_B & \hat{\gamma}_B \end{pmatrix}\begin{pmatrix} r \\ c \\ 1 \end{pmatrix}\right]$$

$$= \begin{pmatrix} g_R' \\ g_B' \end{pmatrix} - \begin{pmatrix} \hat{\alpha}_R' & \hat{\beta}_R' & \hat{\gamma}_R' \\ \hat{\alpha}_B' & \hat{\beta}_B' & \hat{\gamma}_B' \end{pmatrix}\begin{pmatrix} r \\ c \\ 1 \end{pmatrix}. \tag{3-7}$$

It follows that (3-6) may be rewritten as

$$er^2 = \lambda_1 \sum_r \sum_c [\hat{\alpha}_R' r + \hat{\beta}_R' c + \hat{\gamma}_R' - g_R'(r,c)]^2 + \lambda_2 \sum_r \sum_c [\hat{\alpha}_B' r + \hat{\beta}_B' c + \hat{\gamma}_B' - g_B'(r,c)]^2. \tag{3-8}$$

Differentiation of (3-8) with respect to $\hat{\alpha}_R'$, $\hat{\beta}_R'$, and $\hat{\gamma}_R'$, imposition of the conditions $\sum_{r\epsilon R} r = \sum_{c\epsilon C} c = 0$ (corresponding to appropriate choice of origin 0 as indicated in Figure 3-1), and setting the results equal to zero lead to expressions for $\hat{\alpha}_R'$, $\hat{\beta}_R'$, $\hat{\gamma}_R'$, $\hat{\alpha}_B'$, $\hat{\beta}_B'$, and $\hat{\alpha}_B'$ *in terms of their true values* and additive noise terms. The optimal values for $\hat{\alpha}_R'$, $\hat{\beta}_R'$, and $\hat{\gamma}_R'$ are then

$$\begin{aligned} \hat{\alpha}_R' &= \alpha_R' + \sum_r \sum_c r\eta_R'(r,c) / \sum_r \sum_c r^2 \\ \hat{\beta}_R' &= \beta_R' + \sum_r \sum_c c\eta_R'(r,c) / \sum_r \sum_c c^2 \\ \hat{\gamma}_R' &= \gamma_R' + \sum_r \sum_c \eta_R'(r,c) / \sum_r \sum_c 1, \end{aligned} \tag{3-9}$$

where α_R', β_R', γ_R' are the true values of the parameters. If one replaces the subscript R by B, one obtains the analogous relations for the blue parameters. One sees from (3-9) and the corresponding blue relations that the estimates are unbiased. Furthermore, using the covariance matrix for the noise, namely, D_η^{-1}, one may develop the covariance matrix for the estimates. This is seen to be a six by six diagonal matrix with

$$V(\hat{\alpha}_R') = \sigma_1^2 / \sum_r \sum_c r^2$$

$$V(\hat{\beta}_R') = \sigma_1^2 / \sum_r \sum_c c^2 \qquad (3\text{-}10)$$

$$V(\hat{\gamma}_R') = \sigma_1^2 / \sum_r \sum_c 1,$$

just as in Haralick's paper,[6] where $\sigma_1^2 = 1/\lambda_1$. Similarly, for the blue parameters one gets expressions analogous to (3-10) with $\sigma_2^2 = 1/\lambda_2$ replacing σ_1^2. The covariance terms are zero as before.

Now consider any two subrectangles, as, for example, those with local origins O_1 and O_4 in Figure 3-1. Note that these subrectangles meet at O, which is also the midpoint of the line segment joining O_1 to O_4. Assuming a linear model representation over both subrectangles, we would like to test the hypothesis that both linear facets (for red and blue components individually) are really part of the same facet. For this to happen, clearly $\alpha_{1R}' = \alpha_{4R}'$, $\beta_{1R}' = \beta_{4R}'$, $\alpha_{1B}' = \alpha_{4B}'$, and $\beta_{1B}' = \beta_{4B}'$. Now suppose that regions 1 and 4 are equally sized, i.e., have the same number of grid points similarly placed, as indicated in Figure 3-1. Furthermore, suppose that O is not a grid point of either region. For purposes of calculation, then, the two subrectangles are mutually exclusive, and we find that both $N_{1R} = (\hat{\alpha}_{1R}' - \hat{\alpha}_{4R}')(\sum_r \sum_c r^2/2)^{1/2}$ and $N_{2R} = (\hat{\beta}_{1R}' - \hat{\beta}_{4R}')(\sum_r \sum_c c^2/2)^{1/2}$ are normal random variables with mean 0 and variance σ_1^2. Similarly, one observes that their blue counterparts have mean 0 and variance σ_2^2. Now, as in Haralick's paper, we observe that, in order for the sloped surfaces to coincide, it follows that their heights must coincide at point O. The coordinates of O relative to O_1 are seen to be $(\Delta r/2, \Delta c/2)$, and those of O relative to O_4 are $(-\Delta r/2, -\Delta c/2)$, where $(\Delta r, \Delta c)$ represents the position of O_4 relative to O_1. For the red component, the true height of the surfaces at point O is given by $\alpha_{1R}'\Delta r/2 + \beta_{1R}'\Delta c/2 + \gamma_{1R}'$ and $\alpha_{4R}'(-\Delta r/2) + \beta_{4R}'(-\Delta c/2) + \gamma_{4R}'$, respectively. Therefore, under the same sloped surface hypothesis, we must have

$$(\alpha_{1R}' + \alpha_{4R}')\Delta r/2 + (\beta_{1R}' + \beta_{4R}')\Delta c/2 + \gamma_{1R}' - \gamma_{4R}' = 0 \qquad (3\text{-}11)$$

Once (3-11) is true, it is seen that

$$N_{3R} = (\hat{\alpha}_{1R}' + \hat{\alpha}_{4R}')\Delta r/2 + (\hat{\beta}_{1R}' + \hat{\beta}_{4R}')\Delta c/2 + \hat{\gamma}_{1R} - \hat{\gamma}_{4R}' \tag{3-12}$$

is normal with mean 0 and variance

$$\sigma_R^2 = 2\sigma_1^2[(\Delta r/2)^2/\sum_r\sum_c r^2+(\Delta c/2)^2/\sum_r\sum_c c^2+1/\sum_r\sum_c 1]. \tag{3-13}$$

We note also that

$$E[(\hat{\alpha}_{1R}' + \hat{\alpha}_{4R}')(\hat{\alpha}_{1R}' - \hat{\alpha}_{4R}')] = E[\hat{\beta}_{1R}' + \hat{\beta}_{4R}')(\hat{\beta}_{1R}' - \hat{\beta}_{4R}')] = 0.$$

Letting

$$\varepsilon_{iR}^2 = \sum_r\sum_c [\hat{\alpha}_{iR}'r + \hat{\beta}_{iR}'c + \hat{\gamma}_{iR}' - g_R'(r,c)]^2,\ i=1,4, \tag{3-14}$$

and using the results of Haralick's paper, we note that $\varepsilon_{1R}^2/\sigma_1^2$ and $\varepsilon_{4R}^2/\sigma_1^2$ are independent chi-squared random variables, each with $\sum_r\sum_c 1 - 3$ degrees of freedom. Note also that it makes sense for us to consider the red and blue facets individually in the transformed (prime) variables, since the estimators $\hat{\alpha}_R'$, $\hat{\beta}_R'$, $\hat{\gamma}_R'$, $\hat{\alpha}_B'$, $\hat{\beta}_B'$, and $\hat{\gamma}_B'$ are statistically independent. Therefore, one is led to consider the F statistic

$$F_R = \frac{(N_{1R}^2 + N_{2R}^2 + \sigma_1^2 N_{3R}^2/\sigma_R^2)/3}{[(\varepsilon_{1R}^2 + \varepsilon_{4R}^2)/(2\sum_r\sum_c 1 - 6)]} \tag{3-15}$$

and a similar statistic F_B for the blue component. If F_R and F_B are both sufficiently small, we agree to accept the hypothesis that rectangles 1 and 4 are part of the same component.

Now let us adopt a general point of view. Suppose, from practical considerations, that we decide that the "hottest part of a ship" (which might be its engine room) is an item of critical interest as far as identification of the ship is concerned. Let us also agree that hot corresponds to dominance in

amplitude of the blue component over the red component. (In infrared terminology, one might be looking at radiance in one wave band as opposed to another.) We agree to examine closely those pixels for which $g_B(r,c) > T$, where T is a high threshold. Referring to Figure 3-1, let us examine rectangles 1, 2, 3, and 4 in turn. Now simply count the number of pixels n_i for which $g_B > T$, where n_i is the number of such elements in box i. Suppose it turns out that only box 1 contains pixels for which the threshold criterion is valid. Then, if the equal covariance matrix assumption is valid, we apply the analysis of the preceding paragraph to determine whether, in turn, rectangles 1 and 4, 1 and 2, and 1 and 3 can be considered as part of the same sloped surface. We store the parameters for each rectangle in the process, so that we can recover them for later use. If rectangles 1 and 2 are part of the same surface, but that surface is not continuable into boxes 3 and 4, then we have located our region of interest, namely, the union of rectangles 1 and 2, whose boundary is just the horizontal line bisecting the big rectangle. When analyzing regions 1 and 2 to determine whether the same facet applies, we use point P as that point for which the two facets should agree. Similarly, for boxes 1 and 3, we use point Q. Now suppose that the sloped facet for rectangle 1 is not extendible into any of the remaining boxes. If that is the case, we subdivide rectangle 1 into four boxes and apply our threshold logic again to determine those subrectangles containing pixels for which $g_B > T$. Let us record these rectangle numbers. Also, observe that subrectangle of rectangle 1 containing the largest number of pixels for which $g_B > T$. Test that subrectangle against the other subrectangles of box 1 to see whether its sloped facet is extendible. If so, we mark off the appropriate subrectangles and the parameters associated with them and proceed to examine the remaining subrectangles containing pixels for which $g_B > T$. Otherwise, let us further subdivide the given subrectangle and reapply our procedure. Eventually, one of two situations will occur: (a) We find a subbox completely contained within the component of interest, together with other subboxes containing part of the component. (b) We arrive at a box containing too few pixels to apply the sloped facet mode. In case (a), we simply record the subbox of interest, together with its α, β, and γ parameters and proceed to backtrack one step up the quad-tree, looking at the remaining boxes containing the pixels of interest. We use the maximum number criterion again when we dissect the remaining boxes. Case (b) can arise in two circumstances: The

first occurs when every box which could be within the component of interest contains too few pixels. This is a signal that either our pixel structure is too course to resolve what is happening or our threshold criterion is too stringent or both. The second circumstance is that in which we are indeed in the vicinity of a boundary point. In that case, we may mark the box as a "boundary box." We then back up one step in the quad-tree and continue the process to find the remaining boundary points. The entire scheme should yield both the red and blue parameters for the region of interest, together with the boundary boxes. A polygonal line can then be drawn through such boxes, and the resulting line can be accepted as a first-order approximation to the boundary. The logic involved in this process is flow-charted in Figure 3-2.

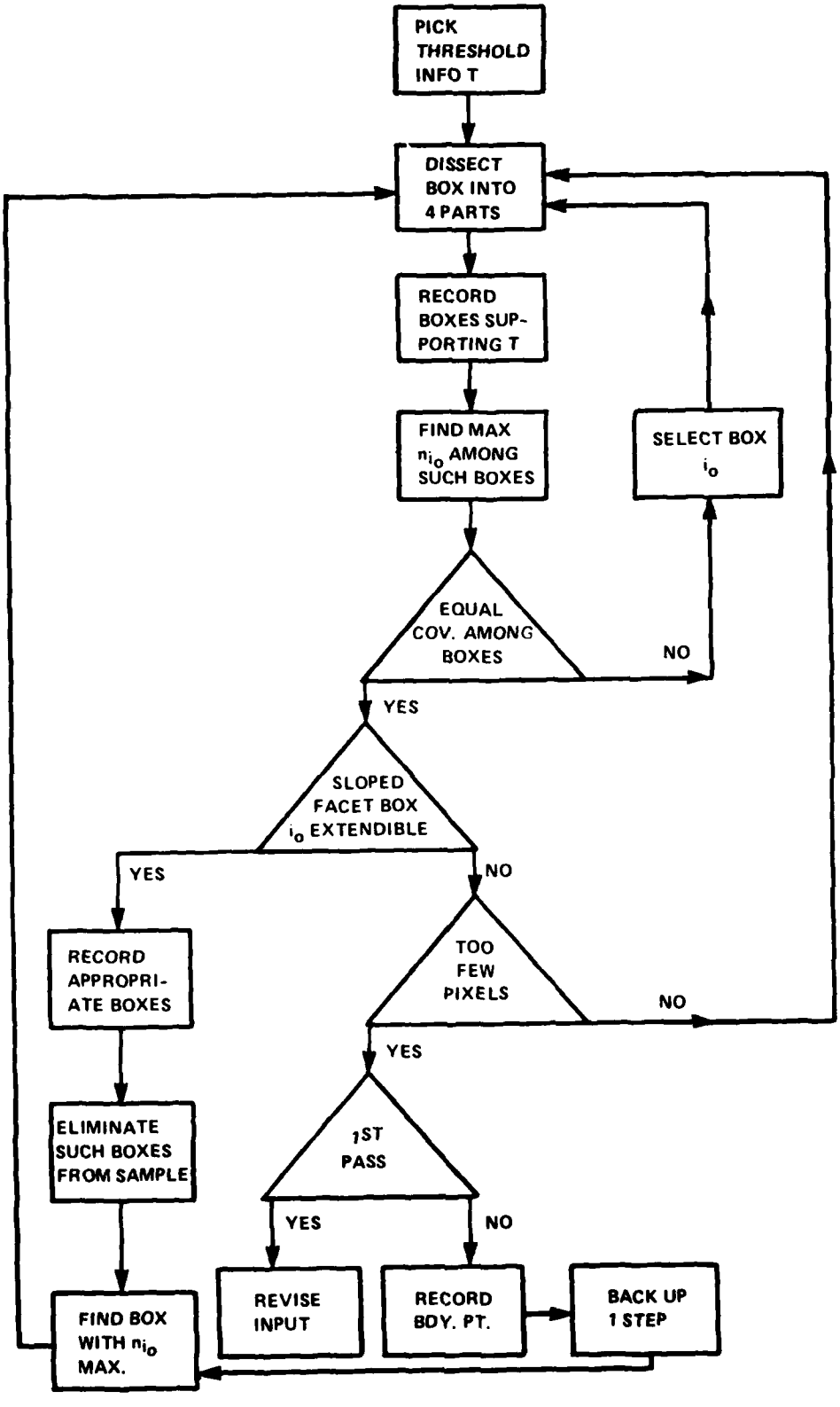


FIGURE 3-2. FLOW CHART FOR QUAD-TREE ASSIGNMENT OF COMPONENT

Now suppose that we have prioritized the subcomponents of the target of interest to us. Let us agree that the blue component intensity is of interest. Scanning the target, one finds that the largest value of $g_B(r,c)$ is T_1. One then looks for the component (or components) containing such pixels. Physically, such components might be the hottest parts of the target. Having found those portions of the image, we eliminate them from consideration and focus our attention on the remaining pixels. Continuing with the same line of reasoning, let us now find those pixels in the rest of the figure for which $g_B(r,c)$ is again maximum, i.e., $g_B(r,c) = T_2$. Pursuing again our quad-tree analysis, we find those subcomponents containing the pixels for which $g_B(r,c) = T_2$. Continuing this process would obviously lead to a first-order dissection of our rectangle into all its primitive subregions. Of course, since g_B is a noisy field, it is not necessarily the case that the components containing those pixels for which g_B is maximum are indeed those corresponding to the hottest portions of the target. That is, the smoothed versions might actually correspond to other subregions. If that is the case, so be it. All that one can say is that the noisy version gives us a priori information and that we shall hopefully learn through a posteriori analysis the underlying nature of the components of interest. If all goes well, there will be a minor reshuffling. Notice that there is a potential saving in this approach in that perhaps only a certain number of subcomponents of a target may be critical to its classification. Also, when all is done, (3-8) should serve as a weighted measure of just how well a sloped facet model of this type represents the data.

THE SOBOLEV MODEL

Before leaving this topic, we shall introduce a new concept which might at times prove useful in generating a sloped facet target description. The idea will be to use the Sobolev norm as either a validation tool or possibly as a means to obtain a better sloped facet model. We must, of course, explain what such a norm is and how it might be helpful. First of all, let us consider the 1-dimensional case. Then the norm is given by

$$||f||_S^2 = \alpha||f||_{L^2}^2 + (1-\alpha)\,||f'||_{L^2}^2 \qquad (3\text{-}16)$$

for any continuously differentiable function f, where

$$||f||^2_{L^2} = \int_a^b f^2(x)\, dx,$$

$0 < \alpha \leq 1$, and the domain of definition of $f(x)$ is the closed interval $[a,b]$. Now a discrete version of (3-16) may be given in which one replaces the notion of derivative by that of divided difference and in which integration is replaced by summation.

Refer then to Figure 3-3. Here the function $f(x)$ is defined only at the

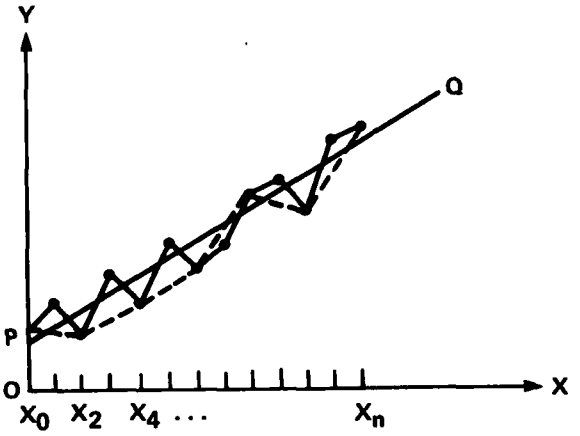


FIGURE 3-3. POLYGONAL LINE APPROXIMATION TO f(x) ON ORIGINAL GRID

control values x_i, $0 \leq i \leq n$, and the polygonal line indicated simply connects all of the points $(x_i, f(x_i))$ in some systematic fashion. As drawn, the average slope of the individual line segments might reasonably approximate that of the line PQ. However, the variance of these slopes from the average may be substantial. If, on the other hand, we double the mesh size, so that we sample only the values of f at x_0, x_2, x_4,..., we see that, roughly speaking, as indicated by the dashed lines, the new set of slopes, half of the previous number, still may have an average value close to the true slope, but now the variance of the slopes seems substantially smaller. Therefore, a rather interesting validation procedure suggests itself. What happens if one decides to use only the values of f(x) on the coarse mesh, together with the slopes (divided differences) as approximations to the derivatives, the slope

informatic eplacing half the function information, and formulate a least squares problem using a Sobolev norm adapted to the mesh? How would the regression line obtained compare with that procured by using all of the function information on the original mesh?

The above is a one-dimensional version of the process in which we are really interested, the latter being two-dimensional. In our situation we have a two-dimensional grid of points. Suppose that one finds, by conventional means, that a sloped facet reasonably represents the noisy data. Could one have done better by doubling or perhaps tripling the mesh size, eliminating half or more of the original grid points, and substituting for the red and blue intensity values divided difference approximations to their first partial derivatives? The new least squares problem would then involve a finite Sobolev norm. The question to be asked is: Would the new technique yield an improvement on the least squares fit? If so, how much of an improvement would there be?

The answers in one dimension were worked out by the author ten years ago.[15] There it was shown that one could derive estimators that were theoretically better than the classical one. It was also shown in another report,[16] in which a shock tube experiment was involved for fitting time of arrival and shock peak pressure data, that such techniques could be used for validation purposes, in that case for determining whether the fit to pressure obtainable through differentiation of a linear fit to time of arrival data was reasonable. We shall here adapt the work in the former report [15] to our 2-D problem. Let us assume that, at each pixel (r,c), we have three measured quantities, written in vector fashion as $(x(r,c), x_r(r,c), x_c(r,c))^t$, where x_r signifies the partial derivative of gray level in the r direction and x_c indicates the partial derivative in the column direction. Of course, e(r,c) itself is to be the gray level at pixel (point) (r,c). In line with our previous development, x would generally be either a red or blue intensity. One would use a divided difference at (r,c) in the r direction to ascertain x_r and, similarly, a divided difference in the c direction to obtain x_c.

There are three types of divided differences which could be used, namely, forward differences, backward differences, and central differences. Inasmuch as there are advantages and disadvantages to the use of the different types, we shall present one model based on the use of forward differences and another

predicated on utilization of central differences. The latter would be more accurate and more straightforward except for the fact, as we shall see, that a coarser grid is demanded. We shall also see that forward differences have the disadvantage of being correlated with function values, whereas central differences are not, provided the points used to calculate them are adequately separated.

Suppose that the pixels are so separated that $x(r,c)$ is independent of $x(r_1,c_1)$ for $(r,c) \neq (r_1,c_1)$. Furthermore, assume that

$$x(r,c) = \alpha r + \beta c + \gamma + n_1(r,c), \tag{3-17}$$

where estimates of α, β, and γ are to be obtained and $n_1(r,c)$ is a zero-mean Gaussian random variable with variance σ^2. We hypothesize that $n_1(r,c)$ has variance σ^2 regardless of the particular point (r,c) involved. Let us treat first the forward divided differences, and suppose that we want to form $x_r(r,c)$, an estimate of the partial derivative in the r direction. We do this with some care so as to insure that $x_r(r,c)$ is independent of $x(r_1,c_1)$, $x_r(r_1,c_1)$, and $x_c(r_1,c_1)$ whenever $(r,c) \neq (r_1,c_1)$. Let us suppose that $x(r,c)$ and $x(r + 2\Delta r,c)$ are two intensity values at adjacent grid points in a column. Assume that the value $x(r + \Delta r,c)$ is available and is statistically independent of $x(r,c)$. Then we shall form the divided difference of $x(r,c)$ and $x(r + \Delta r,c)$ and use that as the estimate for $x_r(r,c)$. Now we find that

$$x_r(r,c) = \frac{x(r + \Delta r,c) - x(r,c)}{\Delta r} = \alpha + \frac{n_1(r + \Delta r,c) - n_1(r,c)}{\Delta r}. \tag{3-18}$$

Two facts emerge immediately from (3-18). First, $x_r(r,c)$ is an unbiased estimator of α. Secondly, the noise term for $x_r(r,c)$ has been derived using (3-17), but is still Gaussian with mean 0. In addition it has variance $2\sigma^2/(\Delta r)^2$. Another observation that we can make is that $x_r(r,c)$ will not be correlated with the vector $(x(r_1,c_1), x_r(r_1,c_1), x_c(r_1,c_1))$, since the latter's noise terms do not appear in (3-18). Using (3-17) and (3-18), one sees that

$$\text{cov}\ (x(r,c),\ x_r(r,c)) = E\left[\frac{n_1(r,c)\ (n_1(r+\Delta r,c) - n_1(r,c))}{\Delta r}\right]$$

$$= -\frac{E\ (n_1^2(r,c))}{\Delta r} \qquad (3\text{-}19)$$

$$= -\sigma^2/\Delta r.$$

Furthermore, the correlation coefficient, defined, as usual, as

$$\rho(x,x_r) = \text{cov}\ (x,x_r)/\sigma_x\sigma_{x_r},$$

becomes

$$\rho(x,x_r) = -1/\sqrt{2}. \qquad (3\text{-}20)$$

Similarly

$$x_c(r,c) = \beta + \frac{n_1(r,c+\Delta c) - n_1(r,c)}{\Delta c}, \qquad (3\text{-}21)$$

representing the partial derivative in the column direction. We assume, as before, that adjacent points in a row are representable by $x(r,c)$ and $x(r,c+2\Delta c)$. One finds that $x_c(r,c)$ is an unbiased estimator of β and that the variance of $x_c(r,c)$ is $2\sigma^2/(\Delta c)^2$. Analogous to (3-19) one has

$$\text{cov}\ (x(r,c),\ x_c(r,c)) = -\sigma^2/\Delta c. \qquad (3\text{-}22)$$

Of course, the correlation coefficient will be the same as in (3-20). Finally, we want cov $(x_r(r,c),\ x_c(r,c))$. Invoking (3-18) and (3-21), we have

$$\text{cov}\ (x_r(r,c),\ x_c(r,c)) = \sigma^2/\Delta r\ \Delta c. \qquad (3\text{-}23)$$

Our model equations may be represented by

$$x(r,c) = \alpha r + \beta c + \gamma + n_1(r,c)$$

$$x_r(r,c) = \alpha + n_2(r,c) \qquad (3\text{-}24)$$

$$x_c(r,c) = \beta + n_3(r,c),$$

where $n_2(r,c)$ is given in eq. (3-18) and $n_3(r,c)$ in eq. (3-21). The covariance matrix for (3-24) is

$$K = \begin{bmatrix} \sigma^2 & -\sigma^2/\Delta r & -\sigma^2/\Delta c \\ -\sigma^2/\Delta r & 2\sigma^2/(\Delta r)^2 & \sigma^2/\Delta r\Delta c \\ -\sigma^2/\Delta c & \sigma^2/\Delta r\Delta c & 2\sigma^2/(\Delta c)^2 \end{bmatrix}$$

$$= \frac{\sigma^2}{a(\Delta r)^2} \begin{bmatrix} a(\Delta r)^2 & -a\Delta r & -\Delta r \\ -a\Delta r & 2a & 1 \\ -\Delta r & 1 & 2/a \end{bmatrix}, \qquad (3\text{-}25)$$

where $a = \Delta c/\Delta r$ is the aspect ratio for the grid. The inverse matrix for (3-25), often called the information matrix, is

$$K^{-1} = \frac{a(\Delta r)^2}{\sigma^2} \begin{bmatrix} 3/a(\Delta r)^2 & 1/a\Delta r & 1/\Delta r \\ 1/a\Delta r & 1/a & 0 \\ 1/\Delta r & 0 & a \end{bmatrix}. \qquad (3\text{-}26)$$

Let us form the likelihood function L[17] for our process:

$$L(\alpha,\beta,\gamma,\sigma^2) = (2\pi)^{-3MN/2}(\det K)^{-MN/2}$$

$$\exp\left[-\frac{1}{2}\sum_{i,j}(n_1(r_i,c_j),n_2(r_i,c_j),n_3(r_i,c_j))\,K^{-1}\right. \qquad (3\text{-}27)$$

$$\left.\cdot(n_1(r_i,c_j),n_2(r_i,c_j)n_3(r_i,c_j))^t\right],$$

where det K is the determinant of K and is $\sigma^6/a^2(\Delta r)^4$, as follows using (3-25). Now maximization of (3-27) is certainly possible, but the calculation of the solution to the likelihood (normal) equations is laborious. It would be computationally advantageous perhaps to approximate the information matrix K^{-1} by a diagonal matrix. It turns out that, for large Δr, the off-diagonal terms in (3-26) may be deleted without appreciable loss. In fact, we shall show that $||K^{-1}-K_1^{-1}||/||K_1^{-1}|| \to 0$ as $\Delta r \to +\infty$, where K_1^{-1} is the diagonal matrix diag $(3/\sigma^2, (\Delta r)^2/\sigma^2, (\Delta c)^2/\sigma^2)$. The double bars denote the norm, here taken to be Euclidean, of the matrix. Using (3-26), we have

$$||K^{-1}-K_1^{-1}||/||K_1^{-1}|| = (2(a^2+1))^{1/2}\Delta r/(9 + (\Delta r)^4 + a^4(\Delta r)^4)^{1/2}, \qquad (3\text{-}28)$$

from which the result follows easily. Also, $\det K_1 = \sigma^6/3a^2(\Delta r)^4$. Let us form this new likelihood function, which we shall call L_1, and consider its negative logarithm. One wants to minimize

$$-\log L_1 = MN\,(3\log 2\pi + \log(\sigma^6/3a^2(\Delta r)^4))/2$$

$$+\sum_{i,j}[3n_1^2(r_i,c_j) + (\Delta r)^2 n_2^2(r_i,c_j) + (\Delta c)^2 n_3^2(r_i,c_j)]/2\sigma^2 \qquad (3\text{-}29)$$

over $(\alpha, \beta, \gamma, \sigma^2)$. In order to do this, we would differentiate (3-29) with respect to α, β, γ and σ^2 in turn. Setting the four quantities thus obtained to zero, we would solve for α, β, γ and σ^2. We suppose that the solution exists and is unique.

The proof that the quantities $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$, $\hat{\sigma}^2$ thus obtained represent the maximizing point of L_1 (or, equivalently, the minimizing point of $-\log L_1$) is not too difficult. In fact, let us first establish the following: For any $V>0$,

there exists an $R > 0$ such that $-\log L_1 > V$ whenever $\alpha^2 + \beta^2 + \gamma^2 + (\sigma^2)^2 > R^2$. Note that

$$-\log L_1 = MN\,(3 \log 2\pi + \log (\sigma^6/3a^2(\Delta r)^4))/2 + \sum_{i,j} [3\,(x_{ij} - \alpha r_i - \beta c_j - \gamma)^2 + (\Delta r)^2(x_{rij} - \alpha)^2 + (\Delta c)^2(x_{cij} - \beta)^2]/2\sigma^2, \tag{3-30}$$

where x_{rij} and x_{cij} are the measured slopes for pixel (r_i, c_j) in the r and c directions, respectively, M is the number of columns, and N is the number of rows. Also, observe that

$$-\sigma^2 \log L_1 = \sigma^2 O(\log(\sigma^2)) + g(\alpha,\beta,\gamma) \equiv f(\sigma^2) + g(\alpha,\beta,\gamma), \tag{3-31}$$

so that $f(\sigma^2) \to +\infty$ as $\sigma^2 \to +\infty$ and $f(\sigma^2) \to 0$ as $\sigma^2 \to 0$. We see that (3-31) is the sum of two quantities, one of which is a function of σ^2 alone and the other of which is a function of α, β, and γ. Furthermore, although $f(\sigma^2)$ becomes negative in $0 < \sigma^2 < 1$, it is bounded there, i.e., there exists a positive constant k such that $|f(\sigma^2)| < k$ whenever σ^2 is in $[0,1]$. Also, it is well-known[18] that there is an $R_1 > 0$ such that $g(\alpha,\beta,\gamma) > V + k$ whenever $\alpha^2 + \beta^2 + \gamma^2 > R_1 - 1$. Therefore, $-\sigma^2 \log L_1 > V$ for $\sigma^2 < 1$ and $\alpha^2 + \beta^2 + \gamma^2 + (\sigma^2)^2 > R_1^2$. Since $\sigma^2 < 1$, $-\log L_1 > V$. Now suppose that $\sigma^2 > 1$, and note that there is a positive number P such that, whenever $(\sigma^2)^2 > P^2$, $f(\sigma^2)/\sigma^2 > V$. In that case, $-\log L_1 > V$ automatically. Finally, assume that $1 < \sigma^2 < P$. Note then, by reasoning similar to that given previously, that there exists an $R_2 > 0$ so that $g(\alpha,\beta,\gamma)/\sigma^2 > g(\alpha,\beta,\gamma)/P > V$ whenever $\alpha^2 + \beta^2 + \gamma^2 > R_2^2 - P^2$. Under these conditions it likewise follows that $-\log L_1 > V$. If we let $R = \max(R_1, R_2)$, we have that $-\log L_1 > V$ whenever $\alpha^2 + \beta^2 + \gamma^2 + (\sigma^2)^2 > R^2$. On the other hand, by continuity, for any $\varepsilon > 0$, there exists a δ_1 sufficiently small and positive such that, when $\alpha^2 + \beta^2 + \gamma^2 + (\sigma^2)^2 < \delta_1^2$, $|-\sigma^2 \log L_1 - g(0,0,0)| < \varepsilon$. That is, on such a closed ball,

$$g(0,0,0) - \varepsilon < -\sigma^2 \log L_1 < g(0,0,0) + \varepsilon \tag{3-32}$$

Assuming that $g(0,0,0) > 0$, one ascertains ε so that $g(0,0,0) - \varepsilon > 0$, also. Now let $\delta_2 = (g(0,0,0) - \varepsilon)/V$, and set $\delta = \min(\delta_1, \delta_2)$. From (3-32)

$$\frac{1}{\sigma^2}(-\sigma^2 \log L_1) > (g(0,0,0) - \varepsilon)/((g(0,0,0) - \varepsilon)/V) = V. \tag{3-33}$$

From (3-33) it follows that $-\log L_1 > V$ on the closed ball $\{(\alpha,\beta,\gamma,\sigma^2) \mid \alpha^2 + \beta^2 + \gamma^2 + (\sigma^2)^2 < \delta^2\}$. If we now choose the compact set $C = \{(\alpha,\beta,\gamma,\sigma^2) \mid \delta^2 < \alpha^2 + \beta^2 + \gamma^2 - (\sigma^2)^2 < R^2\}$, where $\bar{x}_{opt} = (\hat{\alpha},\hat{\beta},\hat{\gamma},\hat{\sigma^2}) \varepsilon C$, it is clear that $\bar{x}_{opt}$ is the minimizing point. Remembering that $\sum r_i = \sum c_j = 0$, the (unique) solution is found to be

$$\hat{\alpha} = (3 \sum_i r_i x_i + (\Delta r)^2 \sum_i \sum_j x_{rij})/M(3 \sum_i r_i^2 + (\Delta r)^2 N) \tag{3-34a}$$

$$\hat{\beta} = (3 \sum_j c_j x_j + (\Delta c)^2 \sum_i \sum_j x_{cij})/N(3 \sum_j c_j^2 + (\Delta c)^2 M) \tag{3-34b}$$

$$\hat{\gamma} = \sum_i x_i/MN \tag{3-34c}$$

$$\hat{\sigma}^2 = \{ \sum_i \sum_j [3(x_{ij} - \alpha r_i - \beta c_j - \gamma)^2 + (\Delta r)^2 (x_{rij} - \alpha)^2 + (\Delta c)^2 (x_{cij} - \beta)^2\}/3MN, \tag{3-34d}$$

where $x_i \equiv \sum_j x_{ij}$ and $x_j \equiv \sum_i x_{ij}$. Here $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are seen to be unbiased estimators of α, β, and γ, respectively.

We next would like to compute the covariance matrix for $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$. This is a straightforward exercise, using the fundamental distributive property cov $(X + Y, Z)$ = cov (X,Z) + cov (Y,Z), where X, Y, and Z are any three random variables and cov means covariance. Again, using the covariance matrix K and the fact that $\sum_j r_i = \sum_j c_j = 0$, one sees that

$$\text{var}(\hat{\alpha}) = \sigma^2(9 \sum_i r_i^2 + 2N(\Delta r)^2)/M\,(3 \sum_i r_i^2 + (\Delta r)^2 N)^2 \tag{3-35a}$$

$$\text{var}(\hat{\beta}) = \sigma^2(9 \sum_j c_j^2 + 2M(\Delta c)^2)/N\,(3 \sum_j c_j^2 + (\Delta c)^2 M)^2 \tag{3-35b}$$

$$\text{var}\ (\hat{\gamma}) = \sigma^2/MN \tag{3-35c}$$

$$\text{cov}\ (\hat{\alpha},\hat{\beta}) = \sigma^2 \Delta r\ \Delta c/(3 \sum_i r_i^2 + (\Delta r)^2 N)\ (3 \sum_j c_j^2 + (\Delta c)^2 M) \tag{3-35d}$$

$$\text{cov}\ (\hat{\alpha},\hat{\gamma}) = -\sigma^2 \Delta r/M\ (3 \sum_i r_i^2 + (\Delta r)^2 N) \tag{3-35e}$$

$$\text{cov}\ (\hat{\beta},\hat{\gamma}) = -\sigma^2 \Delta c/N\ (3 \sum_j c_j^2 + (\Delta c)^2 M). \tag{3-35f}$$

Note that all of the variances and covariances except for var $(\hat{\gamma})$ involve Δr or Δc or both Δr and Δc. Note also that making Δr and Δc larger has the effect of decreasing the values of these parameters. On the other hand, reducing the number of points decreases M and N and the sums of the r_i^2 and c_j^2. This sets up tradeoffs which we shall consider in more detail later.

Let us note that

$$\text{var}\ (\hat{\alpha} r_i + \hat{\beta} c_j + \hat{\gamma}) = r_i^2\ \text{var}\ \hat{\alpha} + c_j^2\ \text{var}\ \hat{\beta} + \text{var}\ \hat{\gamma} + r_i c_j \text{cov}\ (\hat{\alpha},\hat{\beta}) + r_i \text{cov}\ (\hat{\alpha},\hat{\gamma}) + c_j \text{cov}\ (\hat{\beta},\hat{\gamma}) \tag{3-36}$$

for every pair (i,j). Therefore, one can use (3-35 a-f) in order to determine the variance of the estimated facet from the true facet at any point (r_i,c_j).

Now let us go back and consider the real likelihood function L, as given by (3-27). Forming - log L, as we did before, and solving the normal equations, one finds that

$$\hat{\alpha} = \frac{[\sum_{i,j} (3r_i + \Delta r)x_{ij} + \Delta r \sum_{i,j} (r_i + \Delta r)x_{rij} + a\Delta r \sum_{i,j} r_i x_{cij} - \Delta r MN\hat{\gamma}]}{M\,[3 \sum_i r_i^2 + N\,(\Delta r)^2]} \tag{3-37a}$$

$$\hat{\beta} = \frac{[\sum_{i,j} (3c_j + a\Delta r)x_{ij} + \Delta r \sum_{i,j} c_j x_{rij} + a\Delta r \sum_{i,j} (c_j + \Delta r)x_{cij} - a\Delta r\hat{\gamma} MN]}{N\,[3 \sum_j c_j^2 + a^2 M(\Delta r)^2]} \tag{3-37b}$$

$$\hat{\gamma} = \frac{[3 \sum_{i,j} x_{ij} + \Delta r \sum_{i,j} (r_{rij} + a\, x_{cij}) - \Delta r MN\, (\hat{\alpha} + a\hat{\beta})]}{3MN}, \qquad (3\text{-}37c)$$

where $a = \Delta c/\Delta r$, as before. As opposed to (3-34c), (3-37c) contains the influence of the r and c partial derivatives of x_{ij}. Note that $\hat{\alpha}$ and $\hat{\beta}$ depend solely on $\hat{\gamma}$. Indeed, (3-37a) and (3-37b) may be substituted into (3-37c), and one obtains, after a laborious calculation,

$$\hat{\gamma} = \frac{\begin{aligned}&\{3 \sum_{i,j} x_{ij} (3\overline{r_i^2} + (\Delta r)^2)\,(3\overline{c_j^2} + (\Delta c)^2) + \Delta r \sum_{i,j} (x_{rij} - a x_{rij})\\ &\qquad (3\overline{r_i^2} + (\Delta r)^2)(3\overline{c_j^2} + (\Delta c)^2)\\ &-\Delta r (3\overline{c_j^2} + (\Delta c)^2)\,[\sum_{i,j} (3r_i + \Delta r)\, x_{ij} + \Delta r \sum_{i,j} (r_i + \Delta r)\, x_{rij}\\ &\qquad + \Delta c \sum_{i,j} r_i x_{cij}]\\ &-\Delta r (3\overline{r_i^2} + (\Delta r)^2)\,[a \sum_{i,j} (3c_j + \Delta c) x_{ij} + \Delta c \sum_{i,j} c_j x_{rij}\\ &\qquad + a\Delta c \sum_{i,j} (c_j + \Delta r)\, x_{cij}]\}\end{aligned}}{(\Delta c)^2 (\Delta r)^2 + 6\,(\Delta r)^2 \overline{c_j^2} + 27\,\overline{r_i^2}\,\overline{c_j^2} + 6\,\overline{r_i^2} (\Delta c)^2}, \qquad (3\text{-}38)$$

where $\overline{r_i^2} \equiv \sum_i r_i^2/N$ and $\overline{c_j^2} \equiv \sum_j c_i^2/M$. We shall not pursue any further details concerning $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$.

Let us now consider the case where central differences are employed, i.e., instead of (3-18), one uses

$$x_r(r,c) = \frac{x(r+\Delta r,c) - x(r-\Delta r,c)}{2\Delta r} = \alpha + \frac{n_1(r+\Delta r,c) - n_1(r-\Delta r,c)}{2\Delta r}. \quad (3\text{-}39)$$

Then one sees immediately that, provided Δr is so large that the noise sources are uncorrelated, $x(r,c)$ and $x_r(r,c)$ are uncorrelated. Furthermore, we can arrange the grid, in a fashion similiar to that for the one-sided differences, so that $x_r(r,c)$ is not correlated with $x(r_1,c_1)$, $x_r(r_1,c_1)$, and $x_c(r_1,c_1)$ either when $(r,c) \neq (r_1,c_1)$. Clearly $x_r(r,c)$ has mean α and variance $\sigma^2/2(\Delta r)^2$. Likewise one has

$$x_c(r,c) = \beta + \frac{n_1(r,c+\Delta c) - n_1(r,c-\Delta c)}{2\Delta c}, \quad (3\text{-}40)$$

from which one sees that the mean of $x_c(r,c)$ is β and that the variance is $\sigma^2/2(\Delta c)^2$. Adjacent points in a row might be represented by $x(r,c)$ and $x(r,c+4\Delta c)$ in this model and adjacent points in a column by $x(r,c)$ and $x(r+4\Delta r,c)$. Note that this imposes a certain coarseness on the mesh that was not present in the formal difference model. However, we shall see that the central difference model has the advantage of simplicity. The basic model equations are the same as in (3-24). As opposed to (3-25), however, we have a much simpler covariance matrix, namely,

$$K = \begin{bmatrix} \sigma^2 & & 0 \\ & \sigma^2/2(\Delta r)^2 & \\ 0 & & \sigma^2/2(\Delta c)^2 \end{bmatrix} = \frac{\sigma^2}{2(\Delta r)^2} \begin{bmatrix} 2(\Delta r)^2 & & \\ & 1 & \\ & & 1/a^2 \end{bmatrix}. \quad (3\text{-}41)$$

The information matrix K^{-1} is

$$K^{-1} = \frac{2(\Delta r)^2}{\sigma^2} \begin{bmatrix} 1/2(\Delta r)^2 & & 0 \\ & 1 & \\ 0 & & a^2 \end{bmatrix} . \tag{3-42}$$

The negative logarithm of the likelihood function for our process is

$$\begin{aligned} -\log L = MN\,(3\log 2\pi + \log(\sigma^6/4a^2(\Delta r)^4))/2 \\ + \sum_{i,j} [n_1^2(r_i,c_j) + 2(\Delta r)^2 n_2^2(r_i,c_j) + 2(\Delta c)^2 n_3^2(r_i,c_j)]/2\sigma^2, \end{aligned} \tag{3-43}$$

to be minimized over $(\alpha, \beta, \gamma, \sigma^2)$.

Solving the normal equations, we have

$$\hat{\alpha} = \frac{\sum_{i,j} r_i\, x(r_i,c_j) + 2(\Delta r)^2 \sum_{i,j} x_r(r_i,c_j)}{M\,(\sum_i r_i^2 + 2N(\Delta r)^2)} \tag{3-44a}$$

$$\hat{\beta} = \frac{\sum_{i,j} c_j\, x(r_i,c_j) + 2(\Delta c)^2 \sum_{i,j} x_c(r_i,c_j)}{N\,(\sum_j c_j^2 + 2M(\Delta c)^2)} \tag{3-44b}$$

$$\hat{\gamma} = \sum_{i,j} x(r_i,c_j)/MN. \tag{3-44c}$$

One notes that $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are now uncorrelated. Furthermore, the variances of $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are, respectively,

$$\text{var}\,(\hat{\alpha}) = \sigma^2\Big(\sum_i r_i^2 + 2(\Delta r)^2 N\Big)/M\,\Big(\sum_i r_i^2 + 2N(\Delta r)^2\Big)^2 \qquad (3\text{-}45a)$$

$$\text{var}\,(\hat{\beta}) = \sigma^2\Big(\sum_j c_j^2 + 2(\Delta c)^2 M\Big)/N\,\Big(\sum_j c_j^2 + 2M(\Delta c)^2\Big)^2 \qquad (3\text{-}45b)$$

$$\text{var}\,(\hat{\gamma}) = \sigma^2/MN \qquad (3\text{-}45c)$$

The rest of the process parallels that of the earlier calculations, so that we need not discuss that matter further.

Now let us consider the circumstances in which one might use a Sobolev version of the sloped facet model. For illustrative purposes, we examine the forward difference formulation. One might ask the following question: Since making the grid coarser could improve our slope estimates, what happens when we double (or perhaps triple) the grid size, eliminating half (or two thirds) of the gray level information from consideration, but, at the same time, replacing that knowledge with divided difference estimates for the row and column partial derivatives? First of all, note that a least squares fit to intensity data alone corresponds to setting Δr and Δc both to zero in (3-35a-b). This gives

$$\text{var}\,(\hat{\alpha}) = \sigma^2/M \sum_i r_i^2 \qquad (3\text{-}46a)$$

$$\text{var}\,(\hat{\beta}) = \sigma^2/N \sum_i c_j^2. \qquad (3\text{-}46b)$$

Next let us see what happens when we decide to eliminate half the grid points. For purposes of illustration, suppose that both M and N are odd. Then, for example, referring to Figure 3-4, rows might be indexed from top to bottom as shown, where we choose to eliminate from consideration rows r_2, r_4, and r_6. For columns the analysis is similar, proceeding from left to right. In general there will be (N-1)/2 rows and (M-1)/2 columns remaining, and the new grid

spacing in the row direction will be $2\Delta r$, where Δr is the original spacing. Similarly, there is a new column spacing $2\Delta c$, where Δc is the original grid size. It is here that one might use the original grid spacings Δr and Δc to procure divided differences, in line with our previous development. According to (3-35a-f), we obtain a covariance matrix for the estimates $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$, but based on $(\frac{M-1}{2})$ $(\frac{N-1}{2})$ points. The relevant quantities can be calculated from (3-35a-f) after the appropriate substitutions are made. We note that our symmetry conditions $\sum_R r = \sum_C c = 0$ are still valid after the modification process just described, so that the theory is still applicable. It is indeed conceivable that a game could be played to ascertain just how many points could be deleted in order to secure a meaningful reduction in variance. The tradeoffs mentioned before with regard to increase in grid size, corresponding to improvement in slope characteristics, but at the same time corresponding to degradation due to a decrease in number of grid points, come into interesting interplay. Note, from (3-25), that doubling of grid size, leading to a doubling of mesh size for divided difference considerations, reduces variances by a factor of four.

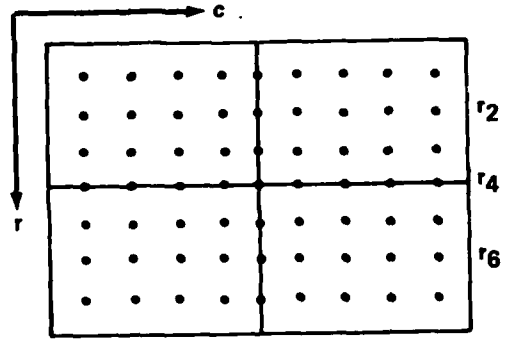


FIGURE 3-4. TESTING THE SOBOLEV FACET MODEL

If one wants to apply Sobolev modeling in a multispectral framework, some preprocessing of information is generally required. For instance, it is quite likely that red and blue intensities, for the same pixel, are correlated. Now suppose that we want to use a linear facet model description for $g_R(r,c)$ and $g_B(r,c)$, where

$$\begin{aligned} g_R(r,c) &= \alpha_R r + \beta_R r + \gamma_R + n_{1R}(r,c) \\ g_B(r,c) &= \alpha_B r + \beta_B r + \gamma_B + n_{1B}(r,c). \end{aligned} \tag{3-47}$$

Having an estimate for the covariance matrix C of the noise vector $(n_{1R},n_{1B})^t$, we know that there exists a 2 by 2 orthogonal matrix P which converts C to a diagonal form, i.e.,

$$PCP^t = D. \tag{3-48}$$

Therefore, applying P to (3-47), we have

$$\begin{aligned} \begin{pmatrix} g_R'(r,c) \\ g_B'(r,c) \end{pmatrix} &= P \begin{pmatrix} g_R(r,c) \\ g_B(r,c) \end{pmatrix} \\ &= P \begin{pmatrix} \alpha_R & \beta_R & \gamma_R \\ \alpha_B & \beta_B & \gamma_B \end{pmatrix} \begin{pmatrix} r \\ c \\ 1 \end{pmatrix} + P \begin{pmatrix} n_{1R}(r,c) \\ n_{1B}(r,c) \end{pmatrix} \\ &= \begin{pmatrix} \alpha_R' & \beta_R' & \gamma_R' \\ \alpha_B' & \beta_B' & \gamma_B' \end{pmatrix} \begin{pmatrix} r \\ c \\ 1 \end{pmatrix} + \begin{pmatrix} n_{1R}'(r,c) \\ n_{1B}'(r,c) \end{pmatrix}, \end{aligned} \tag{3-49}$$

with

$$\operatorname{cov} \begin{pmatrix} n_{1R}'(r,c) \\ n_{1B}'(r,c) \end{pmatrix} = P \operatorname{cov} \begin{pmatrix} n_{1R}(r,c) \\ n_{1B}(r,c) \end{pmatrix} P^t = D, \tag{3-50}$$

so that, in the transformed version, the covariance matrix (and hence the information matrix) is diagonal. Now it is feasible to apply the Sobolev norm to determine estimates $\hat{\alpha}_R'$, $\hat{\beta}_R'$, $\hat{\gamma}_R'$, $\hat{\alpha}_B'$, $\hat{\beta}_B'$, and $\hat{\gamma}_B'$. There will be a 6 by 6 covariance matrix for these quantities expressible as a direct sum of two 3 by 3 covariance matrices, since red and blue information are decoupled from each other. Once these estimates have been determined in primed coordinates, premultiplication of (3-49) by P^t gets one back to the original gray level context, and the variance-covariance properties in the unprimed frame are easily obtained through linear combinations evolving from

$$P^t \begin{pmatrix} \hat{\alpha}_R' & \hat{\beta}_R' & \hat{\gamma}_R' \\ \hat{\alpha}_B' & \hat{\beta}_B' & \hat{\gamma}_B' \end{pmatrix}. \tag{3-51}$$

For example, if

$$P^t = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}, \tag{3-52}$$

then

$$\hat{\alpha}_R = p_{11}\hat{\alpha}_R' + p_{12}\hat{\alpha}_B', \tag{3-53}$$

so that

$$\mathrm{var}(\hat{\alpha}_R) = p_{11}^2 \mathrm{var}\,(\hat{\alpha}_R') + 2p_{11}p_{12}\mathrm{cov}(\hat{\alpha}_R', \hat{\alpha}_B') + p_{12}^2 \mathrm{var}(\hat{\alpha}_B'). \tag{3-54}$$

It must be realized that nothing dictates that we must use divided difference approximations for derivatives. In fact, as was done in one of the author's reports[16], we may use any estimates that we have available in the context of the physical situation.

We conclude this chapter by remarking that what we now have is a first order model for target description in the sense that the mesh is assumed so coarse that we can hypothesize independence of noise from one pixel element to the next and in the sense that no distinction has been made as to the kind of noise present, i.e., instrument noise and process noise (cloud effects, spurious reflections off targets, etc.) have not been distinguished. The separation of such noise sources leads to the conception of Kalman filtering (and other kinds of filtering as well). The next chapter will be devoted to deriving two-dimensional filters in the context of multispectral analysis and to interfacing the sloped facet logic and these filter models.

CHAPTER 4

TWO-DIMENSIONAL FILTERING APPLIED TO COLOR IMAGING

THE KALMAN FILTER

Before embarking on our discussion of the two-dimensional Kalman Filter, let us again adopt a philosophical point of view. Up to this point, we have used the concept of the quad-tree, together with sloped facet modeling and F tests of significance, in order to ascertain components of interest. We have observed that the boundaries of the components flow naturally out of this procedure, there being no need even to make use of edge detection operators based on gradient or gradient-like methods. However, we also noted that inherent in the methodology was the assumption of independence of noise between pixels. This assumption seems to imply a coarse mesh, so that the correlation between points can be neglected. We shall see that a Kalman filter is based on an autoregressive process[19], so that some kind of Markov hypothesis is at work. Therefore, one is able to refine the mesh and take account, in a systematic way, of the noise dependencies thus engendered. Furthermore, we shall be able to separate out the effects of noise entering the process itself from the instrument, or observation, noise. Our idea will be to use the sloped facet model results in two ways: (1) In order to derive an a priori covariance matrix for the state vector so as to implement locally the two-dimensional Kalman filter; (2) in order to generate initial geometrical information about the boundaries of components of interest. Of course, the sloped facet model itself has another virtue: Inasmuch as it provides a first-order description of an image, that description might be good enough. On the other hand, if one is interested in higher levels of detail, where one needs to employ a fine mesh, the Kalman filter should provide an improvement.

The Kalman filter which we now introduce is based on the work of Woods [9] and Kaufman, Woods, Dravida, and Tekalp[8]. It will, however, be generalized to incorporate the aspects of the color imaging procedure. First of all, one should understand the meaning of the state vector as given by Woods. He defines it as the minimum amount of information about past and present estimates needed

to determine an optimal causal estimate of future response given future noisy observations. Generally speaking, the dynamic model for this case is given by

$$\underline{s}(m) = F\underline{s}(m-1) + G\underline{u}(m)$$
$$\underline{r}(m) = H\underline{s}(m) + \underline{v}(m), \tag{4-1}$$

where $\underline{s}$ is the state vector, $\underline{r}$ is the observation vector, $\underline{v}$ is the observation noise vector, and $\underline{u}$ is the process noise vector.

Woods points out that the concept of state vector as given in the preceding paragraph will unfortunately generally lead to a large amount of computation time in situations where the number of pixels in a row, N, is quite a bit larger than the order of the support M of the filter. We shall now define more precisely the state vector and the support M. Underlying the model is the concept of a scalar autoregressive process given by

$$s(m,n) = \sum_{(m-k,n-\ell)\varepsilon R_M(m,n)} c(k,\ell)\, s(m-k,n-\ell) + w(m,n), \tag{4-2}$$

where

$$R_M(m,n) = [(m-k,n-\ell) \mid (1 \leq k \leq M,\ 0 \leq \ell \leq M) \cup (-M \leq k \leq 0,\ 1 \leq \ell \leq M)]. \tag{4-3}$$

The region $R_M(m,n)$ is indicated by the dots in Figure 4-1. Note that

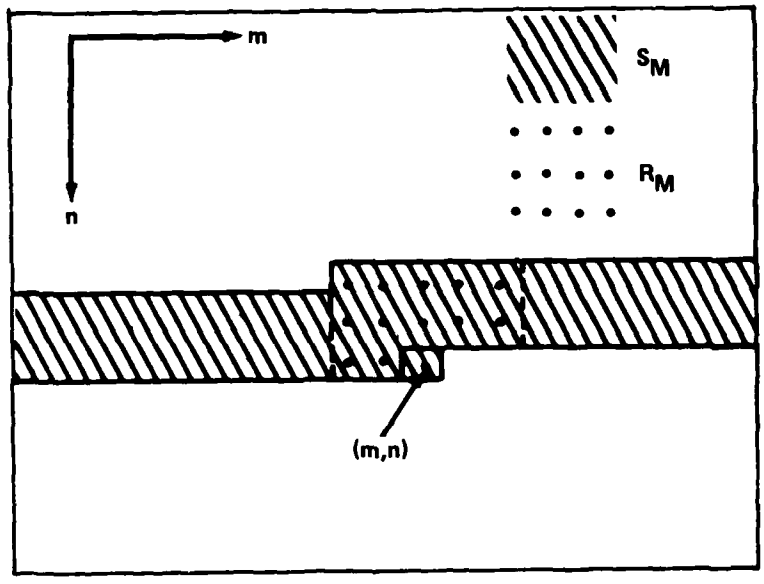


FIGURE 4-1. SUPPORT REGIONS FOR STATE AND PSEUDOSTATE VECTORS

$c(k,\ell)$, the coupling coefficient between $s(m,n)$ and $s(m-k,n-\ell)$, depends only on the relative positions of pixels (m,n) and $(m-k,n-\ell)$. Also, it is assumed that $w(m,n)$ is a zero mean uncorrelated Gaussian two-dimensional field. The state vector is therefore given by

$$\underline{s}(m,n) = [s(m,n),\ s(m-1,n)\ldots,s(1,n);\ s(N,n-1),\ldots,s(1,n-1);$$
$$\ldots;\ s(N,n-M),\ldots,s(m-M,n-M)]^t. \qquad (4\text{-}4)$$

Note that (4-4) contains all relevant information from the past needed to calculate estimates at $(m+1,n),\ldots,(N,n),(1,n+1),(2,n+1),\ldots,(N,n+1),\ldots$, and thus, by definition, is the state vector. Observe, also, that carrying along a certain amount of information as a vector in this way, at the same time, allows us to use the Kalman filter automatically to do a certain amount of smoothing. Another way is, of course, to run a scalar Kalman filter forward and then backward.

Woods introduces the following pseudo-state vector corresponding to a scalar line by line scan:

$$\underline{s}_1(m,n) = [s(m,n),\ldots,s(m-M,n);\ s(m+M,n-1,\ldots,s(m-M,n-1);$$
$$\ldots;\ s(m+M,n-M),\ldots,s(m-M,n-M)]^t. \tag{4-5}$$

In his updating procedure, he chooses only to calculate $\underline{s}_1(m,n)$ at each step rather than $\underline{s}(m,n)$.

Up to this point, we have merely reviewed the work of Woods and his colleagues. In order to adapt their procedures to image modeling, several points should be made: (1) In view of what has transpired in earlier chapters, we see that we may not need to run the Kalman filter over the entire image. It may very well be sufficient to localize it to a subrectangle of interest containing a critical component. (2) If the subrectangle containing our component is "small", it may be unnecessary to resort to the Reduced Update Kalman Filter of Woods. In other words, use of the state vector $\underline{s}(m,n)$ instead of $\underline{s}_1(m,n)$ in the update procedure may not entail that much more work. (3) The sloped facet model should allow us to build a reasonable a priori covariance matrix for our state vector.

In view of the above remarks, we shall not emphasize the Reduced Update Kalman Filter in what follows, but we shall review how one estimates the coefficients $c(k,\ell)$ and the process noise variance Q_w, the latter of which is now generally assumed to vary from point to point. Also, it should be mentioned that the state vector in our case generally consists, as before, of red and blue normalized color components, together possibly with estimates of the first-order partial derivatives. These are, of course, recorded throughout the support region, as reflected in (4-4).

If we are using a primary color representation, possibly together with the first partial derivatives, we must employ a vector analogue of (4-2). For example, one might have

$$g_R(m,n) = \sum_{(m-k,n-\ell)\in R_M(m,n)} c_R(k,\ell)\, g_R(m-k,n-\ell) + w_R(m,n)$$

$$g_B(m,n) = \sum_{(m-k,n-\ell)\in R_M(m,n)} c_B(k,\ell)\, g_B(m-k,n-\ell) + w_B(m,n), \tag{4-6}$$

where one must estimate the coefficients $c_R(k,\ell)$, $c_B(k,\ell)$, together with the noise covariance matrix Q_w for the vector $(w_R, w_B)^t$. We would assume that the vector $(w_R, w_B)^t$ is uncorrelated between pixels, but that w_R and w_B might be correlated at the same pixel. If one has first-order partials for g_R and g_B to be considered, then there would be four more relations analogous to those in (4-6) to be invoked. One would then have a six by six covariance matrix Q_w. As in Woods,[9] one can express the vector $\underline{s}(m,n)$ as $(\underline{s_1}^t(m,n, \underline{s_2}^t(m,n))^t$, with $\underline{s_1}$ given by

$$\begin{aligned}\underline{s_1}(m,n) = [&g_R(m,n),\ g_B(m,n),\ g_R(m-1,n),\ g_B(m-1,n),\ \ldots,\\ &g_R(m-M,n),\ g_B(m-M,n);\ g_R(m+M,n-1),\ g_B(m+M,n-1),\\ &\ldots,\ g_R(m-M,n-1),\ g_B(m-M,n-1);\ \ldots,\\ &g_R(m+M,n-M),\ g_B(m+M,n-M),\ \ldots,\ g_R(m-M,n-M),\\ &g_B(m-M,n-M)]^t.\end{aligned} \tag{4-7}$$

The vector $\underline{s_2}(m,n)$ is then just the remaining part of the state vector $\underline{s}(m,n)$, including both red and blue components. The state dynamical model now becomes

$$\underline{s}(m,n) = C\ \underline{s}(m-1,n) + \underline{w}(m,n), \tag{4-8}$$

where C is the system propagation matrix determined by $\{c_R(k,\ell), c_B(k,\ell)\}$ and by the ordering of the state vector $\underline{s}(m,n)$. The process noise vector is given by

$$\underline{w}(m,n) = (w_R(m,n),\ w_B(m,n),\ 0,\ldots,0)^t. \tag{4-9}$$

There is an observation vector equation

$$\underline{r}(m,n) = H\underline{s}(m,n) + \underline{v}(m,n), \tag{4-10}$$

where $\underline{r}(m,n) = (r_R(m,n),\ r_B(m,n)^t$; $H = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \end{pmatrix}$,

and $\underline{v}(m,n) = (v_R(m,n),\ v_B(m,n))^t$. Now suppose that one has determined, using the quad-tree analysis and sloped facet modeling, a preliminary description of a certain component and its neighborhood in terms of the facet description, as indicated in Figure 4-2. Starting at the upper left-hand corner of the rectangle, one determines, first of all, the facet corresponding to the first pixel. By a facet corresponding to a pixel, we shall mean the following: We identify the

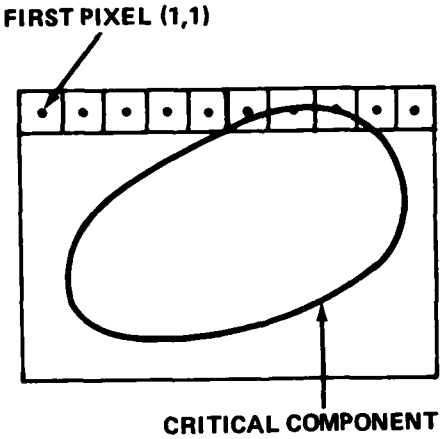


FIGURE 4-2. LOCALIZING A CRITICAL COMPONENT FOR KALMAN FILTERING STUDY

facet by looking at the center of the pixel and the particular facet to which that center pertains. Then the covariance matrix for the first pixel is ascertained simply by computing the sample covariance matrix over all pixels in the domain of that facet (or possibly in some prescribed window about that pixel). For example,

$$\text{var}\,(g_R(1,1)) = \sum_{(i,j)\in D(F_1)} (g_R(i,j) - \hat{\alpha}_R r_i - \hat{\beta}_R c_j - \hat{\gamma}_R)^2/N(F_1), \qquad (4\text{-}11)$$

where $D(F_1)$ is the domain of facet F_1 corresponding to pixel (1,1) and $N(F_1)$ is the number of pixels in the domain of that facet. Also,

$$\text{cov}\,(g_R(1,1),\ g_B(1,1)) = \sum_{(i,j)\in D(F_1)} (g_R(i,j) - \hat{\alpha}_R r_i - \hat{\beta}_R c_j - \hat{\gamma}_R)\,(g_B(i.j) - \hat{\alpha}_B r_i - \hat{\beta}_B c_j - \hat{\gamma}_B)/N(F_1). \qquad (4\text{-}12)$$

Finally, $\text{var}(g_B(1,1))$ is just (4.11) with B replacing R as a subscript. We proceed to march across the first row, computing the same type of information, remembering that, under first order modeling, gray levels are assumed independent from pixel to pixel. In any event, at this level, let us ignore such interpixel correlations. We shall finally procure a covariance matrix P(m,n) for some pixel (m,n) which we shall accept as an initial covariance matrix to start the vector processor. It will be expressed as a direct sum of two by two matrices.

Having thus described how sloped facet modeling might be useful in starting the Kalman filtering model, let us proceed to show how one estimates the coefficients $c_R(k,\ell)$, $c_B(k,\ell)$, and the noise covariance matrix Q_w. First of all, let us assume, as it is customary to do, that one knows the characteristics of his measuring apparatus, so that he can establish the covariance matrix for the measurement noise, namely,

$$Q_v = \begin{pmatrix} \sigma_{vR}^2 & \sigma_{vRB} \\ \sigma_{vRB} & \sigma_{vB}^2 \end{pmatrix}, \qquad (4\text{-}13)$$

where σ_{vR}^2 is the measurement variance for the red component, σ_{vRB} is the measurement covariance for red and blue (which may be zero, for example), and σ_{vB}^2 is the measurement variance for the blue component. Following the development in Kaufman's paper[8], it turns out that the generalization to our situation is straightforward. Let us introduce the matrix J given by

$$J = E\left\{\left[\begin{pmatrix} r_R(m,n) \\ r_B(m,n) \end{pmatrix} - \begin{pmatrix} \underline{c}_R{}^t & 0 \\ 0 & \underline{c}_B{}^t \end{pmatrix} \begin{pmatrix} \underline{r}_R{}^{(1)}(m-1,n) \\ \underline{r}_B{}^{(1)}(m-1,n) \end{pmatrix}\right]\right.$$

$$\left.\left[(r_R{}^t(m,n),\ r_B{}^t(m,n)) - (\underline{r}_R^{(1)}\ (m-1,n)^t,\ \underline{r}_B^{(1)}\ (m-1,n)^t)\begin{pmatrix} \underline{c}_R & 0 \\ 0 & \underline{c}_B \end{pmatrix}\right]\right\} \tag{4-14}$$

with the following elucidation of notation: $r_R(m,n)$ and $r_B(m,n)$ are simply the measured intensities in the red component and blue component at pixel (m,n), respectively. The column vectors of red and blue coupling coefficients are $\underline{c}_R$ and $\underline{c}_B$, as given in equations (4-6). Note that they are defined over the region $R_M(m,n)$, as illustrated in Figure 4-1. The vectors $r_R{}^{(1)}(m-1,n)$ and $r_B{}^{(1)}(m-1,n)$ are pseudovectors given by

$$\begin{aligned} \underline{r}_R{}^{(1)}(m-1,n) = [&r_R(m-1,n),\ldots,r_R(m-M-1,n);r_R(m+M-1,n-1),\\ &\ldots,r_R(m-M-1,n-1);\ldots;r_R(m+M-1,n-M),\\ &\ldots,r_R(m-M-1,n-M)]^t \end{aligned} \tag{4-15}$$

$$\begin{aligned} \underline{r}_B{}^{(1)}(m-1,n) = [&r_B(m-1,n),\ldots,r_B(m-M-1,n);r_B(m+M-1,n-1)\\ &\ldots,r_B(m-M-1,n-1);\ldots;r_B(m+M-1,n-M),\\ &\ldots,r_B(m-M-1,n-M)]^t \end{aligned}$$

Now we also know that

$$\begin{aligned} r_R(m,n) &= g_R(m,n) + v_R(m,n) \\ r_B(m,n) &= g_B(m,n) + v_B(m,n) \end{aligned} \tag{4-16}$$

and that

$$\underline{r}_R^{(1)}(m-1,n) = \underline{g}_R^{(1)}(m-1,n) + \underline{v}_R^{(1)}(m-1,n)$$
$$\underline{r}_B^{(1)}(m-1,n) = \underline{g}_B^{(1)}(m-1,n) + \underline{v}_B^{(1)}(m-1,n), \tag{4-17}$$

where $\underline{g}_R^{(1)}(m-1,n)$ and $\underline{g}_B^{(1)}(m-1,n)$ are obtained from (4-15) upon replacing r by g. Likewise the noise vectors $\underline{v}_R^{(1)}(m-1,n)$ and $\underline{v}_B^{(1)}(m-1,n)$ are found by substituting v for r in (4-15). In (4-16), $v_R(m,n)$ and $v_B(m,n)$ represent the measurement noise at (m,n) in the red and blue components, respectively, and $g_R(m,n)$ and $g_B(m,n)$ are given by (4-6). Substitution of relations (4-16) and (4-17) into (4-14) leads to the fact that

$$J = Q_w + Q_v + \begin{pmatrix} \underline{c}_R^t & 0 \\ 0 & \underline{c}_B^t \end{pmatrix} \tilde{Q}_V \begin{pmatrix} \underline{c}_R & 0 \\ 0 & \underline{c}_B \end{pmatrix}, \tag{4-18}$$

where $\tilde{Q}_v$ has the block form

$$\begin{pmatrix} \sigma_{vR}^2 I & \sigma_{vRB} I \\ \sigma_{vRB} I & \sigma_{vB}^2 I \end{pmatrix}$$

and I is an identity matrix of appropriate size. This matrix is, of course, diagonal when $\sigma_{vRB} = 0$. Once we have estimates for $\underline{c}_R$, $\underline{c}_B$, and J, we can solve (4-18) to obtain Q_w. Following the development in Kaufman's paper, estimation of J, given $\underline{c}_R$ and $\underline{c}_B$, is straightforward. One simply uses a window around pixel (m,n) and computes sample expected values as approximations to true expected values in (4-14).

Let us next discuss the estimation of $\underline{c}_R$ and $\underline{c}_B$. In order to do imbiased image parameter identification, we shall employ observation correlations[20]. Substituting (4-16) and (4-17) into (4-6), we find, in matrix notation, that

$$\begin{pmatrix} r_R(m,n) \\ r_B(m,n) \end{pmatrix} = \begin{pmatrix} \underline{c}_R^{\,t} & 0 \\ 0 & \underline{c}_B^{\,t} \end{pmatrix} \begin{pmatrix} \underline{r}_R^{(1)}(m-1,n) \\ \underline{r}_B^{(1)}(m-1,n) \end{pmatrix} - \begin{pmatrix} \underline{c}_R^{\,t} & 0 \\ 0 & \underline{c}_B^{\,t} \end{pmatrix} \begin{pmatrix} \underline{v}_R^{(1)}(m-1,n) \\ \underline{v}_B^{(1)}(m-1,n) \end{pmatrix} \tag{4-19}$$

$$+ \begin{pmatrix} w_R(m,n) \\ w_B(m,n) \end{pmatrix} + \begin{pmatrix} v_R(m,n) \\ v_B(m,n) \end{pmatrix}.$$

Premultiplying (4-19) by the vector $(r_R(m-k,n-\ell),\ r_B(m-k,n-\ell))$, i.e., taking an inner product, we find

$$\begin{aligned} & r_R(m-k,n-\ell)\ r_R(m,n) + r_B(m-k,n-\ell)\ r_B(m,n) \\ & = r_R(m-k,n-\ell)\ \underline{c}_R^{\,t}\ \underline{r}_R^{(1)}(m-1,n) + r_B(m-k,n-\ell)\ \underline{c}_B^{\,t}\ \underline{r}_B^{(1)}(m-1,n) \\ & \quad - r_R(m-k,n-\ell)\ \underline{c}_R^{\,t}\ \underline{v}_R^{(1)}(m-1,n) - r_B(m-k,n-\ell)\ \underline{c}_B^{\,t}\ \underline{r}_B^{(1)}(m-1,n) \\ & \quad + r_R(m-k,n-\ell)\ w_R(m,n) + r_B(m-k,n-\ell)\ w_B(m,n) \\ & + r_R(m-k,n-\ell)\ v_R(m,n) + r_B(m-k,n-\ell)\ v_B(m,n) \end{aligned} \tag{4-20}$$

Taking expectations of both sides of (4-20) and choosing k and ℓ sufficiently large and positive, one finds that

$$\begin{aligned} & E\,[r_R(m-k,n-\ell)\ r_R(m,n)] + E\,[r_B(m-k,n-\ell)\ r_B(m,n)] \\ & = \underline{c}_R^{\,t}\ E[\underline{r}_R^{(1)}(m-1,n)\ r_R(m-k,n-\ell)] \\ & \quad + \underline{c}_B^{\,t}\ E\,[\underline{r}_B^{(1)}(m-1,n)\ r_B(m-k,n-\ell)]. \end{aligned} \tag{4-21}$$

Then one can obtain estimates for $\underline{c}_R$ and $\underline{c}_B$, namely, $\hat{\underline{c}}_R$ and $\hat{\underline{c}}_B$, by posing a least squares problem in which (4-21) is used. This completes our discussion of the Kalman filtering methodology and of its interconnections with the sloped facet model.

NONLINEAR FILTERING

Linear autoregressive models, together with linear measurement relationships, imply utilization of a linear Kalman filter for smoothing purposes. On the other hand, nonlinear autoregressive models will implicate nonlinear filtering techniques (or perhaps extended Kalman filter algorithms). For example, one might have

$$\begin{aligned} g_R(m,n) &= P_R[\{g_R(m-k,n-\ell)\}] + w_R(m,n) \\ g_B(m,n) &= P_B[\{g_B(m-k,n-\ell)\}] + w_B(m,n), \end{aligned} \tag{4-22}$$

where P_R and P_B are red and blue polynomials, respectively. There may also be some more general functional relationship among the intensities. It will be understood, as before, that $(m-k,n-\ell) \in R_M(m,n)$.

We shall now introduce a completely nonlinear filtering algorithm. It is based on Bayes' theorem and goes as follows: Let us suppose that, at pixel (m,n), one has the state vector $\underline{s}(m,n)$ given in the manner heretofore prescribed with alternating red and blue components. Let $p_b(\underline{s}(m,n))$ be the probability density for $\underline{s}(m,n)$ before updating (i.e., before invoking the measurement at (m,n)). Assume a measurement model of the form

$$\begin{aligned} r_R(m,n) &= h_R(g_R(m,n)) + v_R(m,n) \\ r_B(m,n) &= h_B(g_B(m,n)) + v_B(m,n), \end{aligned} \tag{4-23}$$

where h_R and h_B may be nonlinear functions. Let $\underline{r}(m,n) = (r_R(m,n), r_B(m,n))^t$. We shall assume, as before, that $(v_R,v_B)^t$ is a zero-mean Gaussian vector with covariance matrix Q_v. Suppose, for simplicity, that p_b is also a Gaussian density. Then the probability density $p_a(\underline{s}(m,n)|\underline{r}(m,n))$, namely, the probability of the state vector, given the most recent measurement, is just

$$p_a(\underline{s}(m,n)|\underline{r}(m,n)) = p(\underline{r}(m,n)|\underline{s}(m,n))\, p_b(\underline{s}(m,n))/p(\underline{r}(m,n)). \tag{4-24}$$

From (4-23),

$$p(\underline{r}(m,n)|\underline{s}(m,n)) = \frac{1}{2\pi|Q_v|} e^{-\frac{1}{2}[(v_R,\ v_B)\ Q_v^{-1}\begin{pmatrix}v_R\\ v_B\end{pmatrix}]}. \qquad (4\text{-}25)$$

Also, if p_b is Gaussian, we have

$$p_b(\underline{s}(m,n)) = \frac{1}{(2\pi)^{(NM+M+1)/2}|K|} e^{-\frac{1}{2}(\underline{s}(m,n) - \bar{\underline{s}}(m,n))^t K^{-1}(\underline{s}(m,n) - \bar{\underline{s}}(m,n))}, \qquad (4\text{-}26)$$

where K is the covariance matrix and $\bar{\underline{s}}(m,n)$ is the expected value of $\underline{s}(m,n)$ before updating at (m,n), Finally, $p(\underline{r}(m,n))$ is just the marginal density obtained by integrating the numerator of (4-24) with respect to the state vector $\underline{s}(m,n)$. Note that v_R and v_B in (4-25) are to be replaced by the differences $r_R(m,n) - h_R(g_R(m,n))$ and $r_B(m,n) - h_B(g_B(m,n))$, respectively.

Now that we have $p_a(\underline{g}(m,n)|\underline{r}(m,n))$, how do we advance the process? The clue is to use (4-22) with m replaced by m+1 if we are not at the end of a row and with n replaced by n+1 and m replaced by 1 otherwise. We shall illustrate the process when we are not at the end of a row. In that case, one might set up the system of relations

$$
\begin{aligned}
g_R(m+1,n) &= P_R[\{g_R(m+1-k,n-\ell)\}] + w_R(m+1,n) \\
g_B(m+1,n) &= P_B[\{g_B(m+1-k,n-\ell)\}] + w_B(m+1,n)
\end{aligned}
$$

$$
\begin{aligned}
g_R(m,n) &= g_R(m,n) \\
g_B(m,n) &= g_B(m,n) \\
&\vdots \\
g_R(1,n) &= g_R(1,n) \\
g_B(1,n) &= g_B(1,n) \\
g_R(N,n-1) &= g_R(N,n-1) \\
g_B(N,n-1) &= g_B(N,n-1) \\
&\vdots \\
g_R(1,n-1) &= g_R(1,n-1) \\
g_B(1,n-1) &= g_B(1,n-1) \\
&\vdots \\
g_R(N,n-M) &= g_R(N,n-M) \\
g_B(N,n-M) &= g_B(N,n-M) \\
&\vdots \\
g_R(m+1-M,n-M) &= g_R(m+1-M,n-M) \\
g_B(m+1-M,n-M) &- g_B(m+1-M,n-M),
\end{aligned}
\qquad (4\text{-}27)
$$

representing a transformation from the set $\{w_R(m+1,n),\ w_B(m+1,n),\ \underline{s}_R(m,n)^t,\ \underline{s}_B(m,n)^t\}$ to the set $\{g_R(m+1,n),\ g_B(m+1,n),\ \underline{s}_R(m,n)^t,\ \underline{s}_B(m,n)^t\}$, where

$$
\begin{aligned}
\underline{s}_R(m,n)^t = (&g_R(m,n),\ldots,g_R(1,n),\ g_R(N,n-1),\ldots,g_R(1,n-1),\ldots, \\
&g_R(N,n-M),\ldots,g_R(m+1-M,n-M))
\end{aligned}
\qquad (4\text{-}28)
$$

and $\underline{s}_B(m,n)^t$ is obtained upon replacing R by B in (4-28). The Jacobian of the system (4-27) is just unity. Therefore, the probability density function for the state vector $(\underline{s}_R(m,n)^t,\ \underline{s}_B(m,n)^t,\ w_R(m+1,n),\ w_B(m+1,n))^t$, with $w_R(m+1,n)$ and $w_B(m+1,n)$ replaced by $g_R(m+1,n) - P_R[\{g_R(m+1-k,n-\ell)\}]$ and $g_B(m+1,n) - P_B[\{g_B(m+1-k,n-\ell)\}]$, respectively, will furnish the density at the next point before the measurement update. If (4-24) is then applied with m+1 in place of m, the process will have been successfully advanced. Of course, as the process moves along, the expected state and covariance matrix structure can be computed

using multiple quadrature methods. For example,

$$E(s(i,j)) = \int_{-\infty}^{\infty} \int \ldots \int s(i,j) p_a(\underline{s}(m,n) | r(m,n)) d\underline{s}(m,n), \qquad (4\text{-}29)$$

where $s(i,j)$ is a component of the vector $\underline{s}(m,n)$. Also,

$$E(s(i,j))s(i_1,j_1) = \int_{-\infty}^{\infty} \int \ldots \int s(i,j)s(i_1,j_1) p_a(\underline{s}(m,n) | \underline{r}(m,n)) d\underline{s}(m,n), \qquad (4\text{-}30)$$

where $s(i,j)$ and $s(i_1,j_1)$ are both elements of $\underline{s}(m,n)$. A simplified version of this general process is obtainable by approximating the density of (4-24) by a Gaussian density after computing relations (4-29) and (4-30). A variety of Monte Carlo integration methods may be invoked to perform the integrations in (4-28) and (4-29). Of course, the parameters of the model (4-22), i.e., coefficients like the c_{jk}'s in the linear model and the process noise covariance matrix, would have to be obtained in some fashion, perhaps analogously to the manner in which they were secured in the linear version. This completes our discussion of filtering methodology.

CHAPTER 5
IMAGE BOUNDARY ESTIMATION

Having introduced methodology for smoothing the image, we still have ahead of us the tasks of properly segmenting the target and of identifying the target based on our understanding of the types of components thus obtained and of their positions relative to one another. Of course, in our discussion in Chapter 3, we observed that a first order segmentation of the target would fall out of a region analysis as a natural by-product. In the present chapter, we shall show how geometrical information gained from this first order approach can be input to a more sophisticated procedure for estimating the boundary which makes use of the output of the 2D Kalman filter.

We shall invoke the methodology due to Nahi and Jahanshahi[7]. Their theory is build around the idea of using a replacement process[21] in order to distinguish an object from its background. In their work, a brightness function $b(m,n)$ was defined and conceived of as a combination of an object brightness function $b_o(m,n)$ and a background brightness function $b_b(m,n)$ as follows:

$$b(m,n) = \gamma(m,n)b_{oR}(m,n) + [1-\gamma(m,n)]b_{bR}(m,n), \tag{5-1}$$

where $\gamma(m,n) = 1$ or 0 depending upon whether or not the pixel (m,n) is considered to be a point of the object or a point of the background. Therefore, the real game was that of determining the statistical properties of γ and using them to advantage for object recognition purposes.

Since our model is to be based on the idea of using color imaging, let us introduce the following modification of (5-1):

$$\begin{aligned} b_R(m,n) &= \gamma(m,n)b_{oR}(m,n) + [1-\gamma(m,n)]b_{bR}(m,n) \\ b_B(m,n) &= \gamma(m,n)b_{oB}(m,n) + [1-\gamma(m,n)]b_{bB}(m,n), \end{aligned} \tag{5-2}$$

where the task, as before, will be that of using the statistical properties of γ in order to determine whether or not (m,n) is a point of the object. Again

$\gamma(m,n) = 1$ when (m,n) lies in the object, and $\gamma(m,n) = 0$ when (m,n) is outside the object. In addition, background likewise generally has color characteristics. Therefore, we should observe some fundamental level of red and blue in the background as well.

As was done in the Kalman filtering analysis, it is convenient for processing purposes to think in terms of row by row (or perhaps column by column) scanning. In the Nahi-Jahanshahi paper, it was assumed that the object whose boundary one would like to ascertain is horizontally (or possibly vertically) convex. An object $E \subseteq R^2$ is said to be horizontally convex if, for $\underline{x}^{(1)} = (x_1^1, x_2^1) \in E, \underline{x}^{(2)} = (x_1^2, x_2^2) \in E$, with $x_1^1 \neq x_1^2$ and $x_2^1 = x_2^2, \alpha\underline{x}^{(1)} + (1-\alpha)\underline{x}^{(2)} \in E$, $0 < \alpha < 1$. An example of a horizontally convex object is given in Figure 5-1. However, this hypothesis may very well be over-restrictive. For example, a

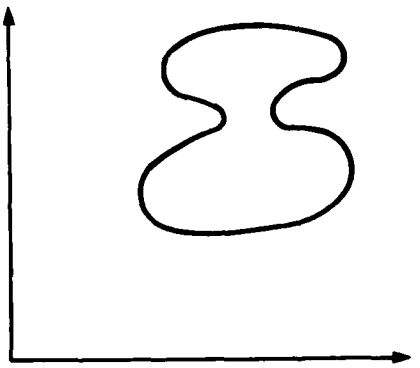


FIGURE 5-1. A HORIZONTALLY CONVEX OBJECT

character like the letter N shown in Figure 5-2 is neither horizontally nor vertically convex. We would like to be able to treat such objects in the present context. Also, objects with holes should be treatable. Therefore, we shall provide a generalization of the Nahi-Jahanshahi theory in this regard.

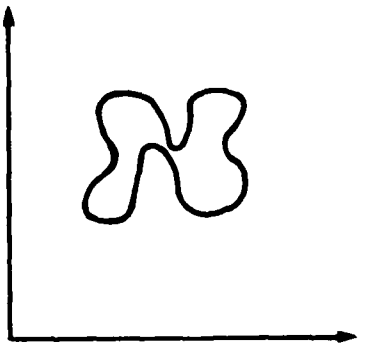


FIGURE 5-2. A NON-HORIZONTALLY CONVEX SET

We shall assume that we have used sloped facet modeling in order to isolate a component of interest, together with a first-order approximation to its boundary. Suppose next that we have localized the component to a rectangle (perhaps one in the quad-tree analysis described in Chapter 3). Our task will be to describe the use of a vector filter processor in conjunction with a boundary estimation scheme similar to that of Nahi and Jahanshahi in order properly to estimate $\gamma(m,n)$. The replacement concept is again to be invoked, whereby we run the Kalman or nonlinear filter processor for both the background and the object individually. The scanner output is, in scalar form,

$$s_R(k) = \lambda(k)\, s_{oR}(k) + [1-\lambda(k)]\, s_{bR}(k)$$
$$s_B(k) = \lambda(k)\, s_{oB}(k) + [1-\lambda(k)]\, s_{bB}(k), \qquad (5\text{-}3)$$

where $s_{oR}(k)$ is the object red component gray level at pixel k and $s_{bR}(k)$ is the background red component at point k. Similar remarks can be made about the blue component. We shall use a number of lines of output from the vector processor at most equal to the order M of the filter; and, as we execute the filter, we

shall attempt dynamically to produce a Nahi-Jahanshahi type boundary estimation procedure.

Now let us proceed with the details of the estimation. The statistics of the process $\lambda(k)$ are to be obtained. We assume that we have before us a first order approximation to the boundary as given through a sloped facet approach. The geometrical information thus provided is to be input, together with the Kalman or nonlinear filter results, in determining a better boundary estimate. In the Nahi-Jahanshahi formulation, the problem was to determine, for each scan line r, a vector $\underline{w}_r = (\alpha_r, \beta_r)$, where α_r is the first boundary element encountered by the scanner and β_r the second boundary element encountered by the scanner. This representation is, of course, somewhat restrictive, as it incorporates the concept of horizontal convexity. Let us suppose, as an alternative, that, based on our first order boundary description, we agree to subdivide the rectangle containing the component of interest into a number of horizontal sections in each of which we have a given number n_p of boundary points per scan line, $1 \leq p \leq NS$. Here NS is the number of sections. Furthermore, let us assume that the number of lines in each section p does not exceed the order M of the autoregressive process involved. Our idea is to use the same type of methodology proposed in the Nahi-Jahanshahi paper, but to do so to better advantage by appealing to a more thorough preliminary geometrical description. To establish notation, let m_{1p} be the first scan line of section p and m_{2p} the last scan line. Also, let $\underline{w}_{rp} = (\alpha_{1rp}, \alpha_{2rp}, \ldots, \alpha_{n_p rp})$ represent

the vector of boundary points in, for example, a left-right scan of the rth line of section p. We shall run through the entire process, first of all, for section 1. Let us be clear concerning the statistical model to be used. Suppose that we base it on a linear autoregressive model such as that given in (4-6). Then one has a state dynamical system afforded by (4-8) with the measurement model (4-10). In particular, consider equation (4-4) with the pair (m,n) replaced by (M+1,M+1). Generalizing to the colored picture representation, one has

$$\underline{s}(M+1,M+1) = [g_R(M+1,M+1),\ g_B(M+1,M+1),\ g_R(M,M+1),\ g_B(M,M+1),$$

$$\ldots,\ g_R(1,M+1),\ g_B(1,M+1);\ g_R(N,M),\ g_B(N,M),\ \ldots,\ g_R(1,M),\ g_B(1,M);$$

$$\ldots;\ g_R(N,1),\ g_B(N,1),\ \ldots,\ g_R(1,1),\ g_B(1,1)]^t. \tag{5-4}$$

Note that $\underline{s}_d(M+1,M+1)$, which we define to be $\underline{s}(M+1,M+1)$ stripped of its first $2(M+1)$ elements, represents the state of the first M rows. Equation (4-8) with (m,n) replaced by (M+1,M+1) is, of course,

$$\underline{s}(M+1,M+1) = C\underline{s}(M,M+1) + \underline{w}(M+1,M+1) \tag{5-5}$$

Forming a vector from the 2-vectors $\underline{r}(M+1,M+1)$, $\underline{r}(M,M+1)$, ..., $\underline{r}(1,1)$, we have

$$\underline{r}_1(M+1,M+1) \equiv (\underline{r}^t(M+1,M+1),\ \underline{r}^t(M,M+1),\ \ldots,\ \underline{r}^t(1,1)^t. \tag{5-6}$$

The measurement equation which we now invoke is

$$\underline{r}_1(M+1,M+1) = \underline{s}(M+1,M+1) + \underline{v}_1(M+1,M+1), \tag{5-7}$$

where

$$\underline{v}_1(M+1,M+1) \equiv (\underline{v}^t(M+1,M+1),\ \underline{v}^t(M,M+1),\ \ldots,\ \underline{v}^t(1,1))^t. \tag{5-8}$$

We then run the model (5-5), together with (5-7), for both the background and the object. That is, we prepare to execute the replacement process. Let $\bar{\underline{s}}_+(M,M+1)$ be the expected state vector after the measurement update at pixel (M,M+1) and $\bar{\underline{s}}_-(M+1,M+1)$ the expected state vector at (M+1,M+1) <u>before the update</u>. Since the estimate before updating is obtained by using the transition matrix C to propagate the state, one has $\bar{\underline{s}}_-(M+1,M+1) = C\bar{\underline{s}}_+(M,M+1)$. Consider then the equation

$$\begin{aligned}\underline{r}_1(M+1,M+1) &= C\bar{\underline{s}}_+(M,M+1) + \underline{v}_2(M+1,M+1) \\ &= \bar{\underline{s}}_-(M+1,M+1) + \underline{v}_2(M+1,M+1).\end{aligned} \tag{5-9}$$

One may check, in a straightforward manner, that $\underline{v}_2$ has mean zero. It remains to determine the covariance matrix for $\underline{v}_2$. One has

$$
\begin{aligned}
E(\underline{v}_2(M+1,M+1)\underline{v}_2{}^t(M+1,M+1)) &= E[(\underline{r}_1(M+1,M+1)-\bar{\underline{s}}_-(M+1,M+1))(\underline{r}_1(M+1,M+1)-\bar{\underline{s}}_-(M+1,M+1))^t] \\
&= E[(\underline{v}_1(M+1,M+1) + \underline{s}(M+1,M+1) - \bar{\underline{s}}_-(M+1,M+1)) \\
&\quad (\underline{v}_1(M+1,M+1) + \underline{s}(M+1,M+1) - \bar{\underline{s}}_-(M+1,M+1))^t] \\
&= E[(\underline{v}_1(M+1,M+1) + \tilde{\underline{s}}_-(M+1,M+1)) \\
&\quad .(\underline{v}_1(M+1,M+1) + \tilde{\underline{s}}_-(M+1,M+1))^t] \\
&= Q_{v_1} + P_-(M+1,M+1),
\end{aligned}
\tag{5-10}
$$

where $\tilde{\underline{s}}_- \equiv \underline{s} - \bar{\underline{s}}_-$.

It is seen that the term $E(\underline{v}_1(M+1,M+1)\ \tilde{\underline{s}}_-(M+1,M+1)^t)$ involved in (5-10) is zero, since $\tilde{\underline{s}}_-(M+1,M+1)$ does not incorporate the measurement update at (M+1,M+1). $P_-(M+1,M+1)$ is, by definition, the state covariance matrix at (M+1,M+1) before updating. Now let us strip off the first 2(M+1) elements of each of the three vectors appearing in (5.9). This leads to an equation

$$
\underline{r}_{1d}(M+1,M+1) = \bar{\underline{s}}_{-d}(M+1,M+1) + \underline{v}_{2d}(M+1,M+1), \tag{5-11}
$$

where d represents the deletion. The covariance matrix for $\underline{v}_{2d}$ is, of course, just the appropriate principal submatrix of (5-10). Let $Q_{\underline{v}_{2d}}$ be that covariance matrix, and suppose that we find an orthogonal matrix A which diagonalizes $Q_{\underline{v}_{2d}}$. That is, we have

$$
D(\underline{\lambda}) = \text{diag}\ (\lambda_1,\lambda_2,\ldots) = AQ_{\underline{v}_{2d}}A^T. \tag{5-12}
$$

If we pre- and post-multiply by

$$D_1(\lambda) = \text{diag}\,(\lambda_1^{-\frac{1}{2}}, \lambda_2^{-\frac{1}{2}}, \ldots),$$

we have

$$I = D_1(\lambda) A Q_{\underline{v}_{2d}} A^t D_1(\lambda). \tag{5-13}$$

Thus, we may premultiply (5-11) by $D_1(\lambda)A$ and obtain

$$\underline{r}_{1d}'(M+1,M+1) = \underline{\bar{s}}_{-d}'(M+1,M+1) + \underline{v}_{2d}'(M+1,M+1), \tag{5-14}$$

where now $\underline{v}_{2d}'(M+1,M+1)$ has the identity matrix as covariance matrix.

Equation (5-14), in conjunction with the replacement concept, may be construed in the following way: After executing the sloped facet model, suppose that we have defined an object and its boundary to first order accuracy and have localized the object to a rectangle. Realizing that what we want to do is to perturb the boundary using the replacement concept, we agree to use the covariance matrix $Q_{\underline{v}_{2d}}$ above and to apply the matrix $D_1(\lambda)A$ in order to transform the state vector. We map the background and object intensities by $D_1(\lambda)A$, and we apply the replacement principle to the transformed quantities. The important point is that the noise covariance is provided by (5-10) as applied to a first order description of the object or background, respectively. Such covariance may be computed without exercising the filter. The filter itself is to be executed for both the background and the object processes in the case where either fills the entire rectangle. The object process is to be construed as that pertaining to the actual scene consisting of object as influenced by background and background as influenced by object. The background process, on the other hand, is that obtained after subtracting an object's influence. An example is that of a ship subject to environmental effects. Taking out the ship, one has only background information. The ship plus its background constitutes the object process. Now there are two kinds of boundaries with which we are concerned: external and internal. For an external boundary, such as that serving to outline a ship, there is no problem with employing the usual replacement process. However, once internal boundaries are

considered, one has to modify his notion of background. In effect, if we know the boundary of the component of interest, then we would want to extrapolate the gray scale values for the rest of the target across that boundary. Then our new background would consist of the original background external to the target, the gray scale values of the target itself minus the component, and certain artificial values created by extrapolation into the "removed component." Now a Kalman filter, in conjunction with a reasonable first-order description of the boundary, can be used to "create the background." To do this, one may conceive of the removal of the component as synonymous with the deletion of data. Therefore, one needs to run a Kalman filter (or other filter) over the object in which data inside the first order boundary estimate mentioned above have infinite (very large) measurement and process variance. This procedure will have the desired effect of propagating the gray scale estimates into the gap created by the removed part. In this way one could conceivably improve on one's understanding of internal boundaries. In the following discussion, we shall give the details of the boundary estimation method.

We are now in a position to pursue the analysis further. From the above remarks, we shall, without loss of generality, assume a model of the form

$$\underline{Y} = \underline{S}(M,\underline{W}) + \underline{U}, \tag{5-15}$$

where $\underline{Y}$ is the set of measurements of both red and blue intensities for the first M rows, $\underline{S}$ is the state vector which starts at row m_{11} and ends at row m_{21}, there being M or fewer rows altogether. Now remember that the first section of the NS sections is characterized by the presence of n_1 boundary points, where $\underline{w}_{r1}$ denotes the vector of boundary points in the rth row of that slice. Generally speaking, n will be an even integer; but, when two boundary points coalesce, a special case arises in which the number of distinct points is odd. $\underline{W}$ is a vector formed from $\underline{w}_{11},\underline{w}_{21},\ldots$, up to the number of rows in the section, and $\underline{U}$ is the noise term whose covariance matrix is the identity matrix.

We want to be a bit more specific about the vector $\underline{S}(M,\underline{W})$, since we need to distinguish between points within the component of interest and those outside of it. In conformance with the notation in the Nahi-Jahanshahi paper, let $s_{bR}(k)$ and $s_{bB}(k)$ represent the red and blue intensity, respectively, at a point outside the component, i.e., in the background. Also, let $s_{OR}(k)$ and $s_{OB}(k)$ be

the analogous quantities at object (component) points. As in the above quoted paper, let J be the number of columns of pixels. Then, for the first m_{11}-1 rows (m_{11} being thought of as random), all the s values are background quantities; whereas, for $m_{11} < r \leq m_{21}$, one has the following row pattern:

$$
\begin{aligned}
&s_{bR}[(r-1)J+1], \; s_{bB}[(r-1)J+1], \; \ldots, \; s_{bR}[(r-1)J+\alpha_{1r}-1], \\
&s_{bB}[(r-1)J+\alpha_{1r}-1], \; s_{oR}[(r-1)J+\alpha_{1r}], \; s_{oB}[(r-1)J+\alpha_{1r}], \; \ldots, \\
&s_{oR}[(r-1)J+\alpha_{2r}], \; s_{oB}[(r-1)J+\alpha_{2r}), \; s_{bR}[(r-1)J+\alpha_{2r}+1], \\
&s_{bB}[(r-1)J+\alpha_{2r}+1], \; \ldots, \; s_{bR}[(r-1)J+\alpha_{3r}-1], \; s_{bB}[(r-1)J+\alpha_{3r}-1], \\
&s_{oR}[(r-1)J+\alpha_{3r}], \; s_{oB}[(r-1J+\alpha_{3r}], \; \ldots, \; s_{oR}[(r-1)J+\alpha_{n_1 r}], \\
&s_{oB}[(r-1)J+\alpha_{n_1 r}], \; s_{bR}[(r-1)J+\alpha_{n_1 r}+1], \; s_{bB}[(r-1)J+\alpha_{n_1 r}+1], \\
&\ldots, \; s_{bR}(rJ), \; s_{bB}(rJ).
\end{aligned}
\tag{5-16}
$$

Note that, in our formulation, m_{21} is generally not random. We shall attempt to maximize the joint probability density function $p(\underline{Y}, \underline{M}, \underline{W})$, which is just

$$
p(\underline{Y},\underline{M},\underline{W}) = p(\underline{Y}|\underline{M},\underline{W}) \; p(\underline{W}|\underline{M}) \; p(\underline{M}). \tag{5-17}
$$

If $\underline{U}$ is assumed to be Gaussian, we have, therefore,

$$
\begin{aligned}
p(\underline{Y},\underline{M},\underline{W}) = (2\pi)^{-N/2} \exp \{ &- \tfrac{1}{2} [\underline{Y} - \underline{S}(\underline{M},\underline{W})]'[\underline{Y} - \underline{S}(\underline{M},\underline{W})] \\
&+ \ln p(\underline{W}|\underline{M}) + \ln p(\underline{M})\},
\end{aligned}
\tag{5-18}
$$

where $N = 2Jm_{21}$. Maximizing (5-18) with respect to $\underline{M}$ and $\underline{W}$ is equivalent to

$$\max_{\underline{M},\underline{W}}\{-\frac{1}{2}[\underline{Y} - \underline{S}(\underline{M},\underline{W})]'[\underline{Y} - \underline{S}(\underline{M},\underline{W})] + \ln p(\underline{W}|\underline{M}) + \ln p(\underline{M})\}, \tag{5-19}$$

or

$$\min_{\underline{M},\underline{W}} \{\underline{S}'(\underline{M},\underline{W})\,[\underline{S}(\underline{M},\underline{W}) - 2\underline{Y}] - 2 \ln p(\underline{W}|\underline{M}) - 2 \ln p(\underline{M})\}. \tag{5-20}$$

Let us transform the first of the three terms within the curly brackets of (5-20) into scalar notation. We have then

$$\underline{S}'[\underline{S}(\underline{M},\underline{W}) - 2\underline{Y}]$$

$$= \sum_{r=1}^{m_{11}-1} \sum_{k=(r-1)J+1}^{rJ} K_b(k)$$

$$+ \sum_{r=m_{11}}^{m_{21}} \left[\sum_{k=(r-1)J+1}^{(r-1)J+\alpha_{1r}-1} K_b(k) + \sum_{k=(r-1)J+\alpha_{1r}}^{(r-1)J+\alpha_{2r}} K_o(k) \right. \tag{5-21}$$

$$+ \ldots + \sum_{k=(r-1)J+\alpha_{n_1-1,r}}^{(r-1)J+\alpha_{n_1 r}} K_o(k)$$

$$\left. + \sum_{k=(r-1)J+\alpha_{n_1 r}+1}^{rJ} K_b(k) \right],$$

where

$$K_b(k) = K_{bR}(k) + K_{bB}(k)$$
$$K_o(k) = K_{oR}(k) + K_{oB}(k), \tag{5-22}$$

with

$$K_{bR}(k) = s_{bR}(k)\,[s_{bR}(k) - 2y_R(k)]$$
$$K_{oR}(k) = s_{oR}(k)\,[s_{oR}(k) - 2y_R(k)] \tag{5-23}$$

and relations similar to (5-23) for the blue components. Note that the summands in (5-21) are all known, the unknowns being the α's and m_{11}.

Now let us add to and subtract from (5-21) the following:

$$\sum_{r=m_{11}}^{m_{21}} \Bigg[\sum_{k=(r-1)J+\alpha_{1r}}^{(r-1)J+\alpha_{2r}} K_b(k) + \sum_{k=(r-1)J+\alpha_{3r}}^{(r-1)J+\alpha_{4r}} K_b(k) + \ldots + \sum_{k=(r-1)J+\alpha_{n_1-1,r}}^{(r-1)J+\alpha_{n_1 r}} K_b(k) \Bigg] \tag{5-24}$$

(5-19) is thus converted to

$$\begin{aligned} \underline{S}'(\underline{M},\underline{W})\,[\underline{S}(\underline{M},\underline{W}) - 2\underline{Y}] &= \sum_{r=1}^{m_{21}} \sum_{k=(r-1)J+1}^{rJ} K_b(k) \\ &+ \sum_{r=m_{11}}^{m_{21}} \Bigg[\sum_{k=(r-1)J+\alpha_{1r}}^{(r-1)J+\alpha_{2r}} [K_o(k) - K_b(k)] \\ &+ \sum_{k=(r-1)J+\alpha_{3r}}^{(r-1)J+\alpha_{4r}} [K_o(k) - K_b(k)] \\ &+ \ldots + \sum_{k=(r-1)J+\alpha_{n_1-1,r}}^{(r-1)J+\alpha_{n_1 r}} [K_o(k) - K_b(k)] \Bigg]. \end{aligned} \tag{5-25}$$

The first group of terms in (5-25) does not depend on m_{11} or the α's, so they may be eliminated from consideration in (5-20). In other words, remembering that n_1 is generally a positive, even integer and letting

$$T(\underline{w}_r) = \sum_{j=1}^{n_1/2} \sum_{k=(r-1)J+\alpha_{2j-1,r}}^{(r-1)J+\alpha_{2j,r}} [K_o(k) - K_b(b)],$$

one wants to consider the minimization problem

$$\min_{\underline{M},\underline{W}} [-2\ln p(\underline{M}) - 2\ln p(\underline{W}|\underline{M}) + \sum_{r=m_{11}}^{m_{21}} T(\underline{w}_r)]. \tag{5-26}$$

Now let us examine the nature of $p(\underline{M})$ and $p(\underline{W}|\underline{M})$. First of all, $p(\underline{M})$ is just $p(m_{11})$ for section 1, since m_{21} is preassigned. This reflects the fact that we do not know with certainty the first row of the component. Also, in general, let us suppose that boundary point α_{srp}, namely, boundary point s in row r of section p, is conditioned only on $\alpha_{s-1,r,p}$ and $\alpha_{s,r-1,p}$, its nearest boundary points "preceding" it in row r and r-1, respectively. It follows that

$$p(\underline{W}|\underline{M}) = p(\underline{w}_{m_{21}}|\underline{w}_{m_{21}-1})\, p(\underline{w}_{m_{21}-1}|\underline{w}_{m_{21}-2}) \cdots p(\underline{w}_{m_{11}+1}|\underline{w}_{m_{11}})\, p(\underline{w}_{m_{11}}), \tag{5-27}$$

where, generally,

$$p(\underline{w}_r|\underline{w}_{r-1}) = p(\alpha_{1r},\alpha_{2r},\cdots,\alpha_{n_1 r}|\alpha_{1,r-1},\alpha_{2,r-1}\cdots,\alpha_{n_1,r-1})$$

$$= p(\alpha_{n_1 r}|\alpha_{n_1-1,r},\alpha_{n_1,r-1}) \cdot p(\alpha_{n_1-1,r}|\alpha_{n_1-2,r},\alpha_{n_1-1,r-1})$$

$$\cdot\ p(\alpha_{n_1-2,r}|\alpha_{n_1-3,r},\ \alpha_{n_1-2,r-1}) \cdots \tag{5-28}$$

$$\cdot\ p(\alpha_{1r}|\alpha_{1,r-1}).$$

From (5-27) one has

$$\ln p(\underline{W}|\underline{M}) = \sum_{r=m_{11}}^{m_{21}} \ln p(\underline{w}_r|\underline{w}_{r-1}), \tag{5-29}$$

where $p(\underline{w}_{m_{11}}|\underline{w}_{m_{11}-1}) \equiv p(\underline{w}_{m_{11}})$. Substitution of (5-29) into (5-26) yields

$$\min_{\underline{M},\underline{W}} \{-2 \ln p(m_{11}) + \sum_{r=m_{11}}^{m_{21}} [T(\underline{w}_r) - 2 \ln p(\underline{w}_r|\underline{w}_{r-1})]\} \tag{5-30}$$

The reader will, of course, observe that what we have done so far is to use the Nahi-Jahanshahi analysis in a slightly more general format. We shall continue to follow the general outline of their modeling procedure. Therefore, let us fix m_{11} in (5-29) equal to that dictated by the sloped facet model result. Then all we need do is minimize with respect to $\underline{W}$ the sum of the terms from m_{11} to m_{21} indicated in (5-30) Let us consider, first of all, the natural logarithm of (5-28). We have

$$\ln p(\underline{w}_r|\underline{w}_{r-1}) = \sum_{i=1}^{n_1} \ln p(\alpha_{ir}|\alpha_{i-1,r},\ \alpha_{i,r-1}), \tag{5-31}$$

so that the minimization problem which we would like to consider is

$$\min_{\underline{W}} \{ \sum_{r=m_{11}}^{m_{21}} [T(\underline{w}_r) - 2 \sum_{i=1}^{n_1} \ln p(\alpha_{ir}|\alpha_{i-1,r},\ \alpha_{i,r-1})]\}. \tag{5-32}$$

The problem is that of providing a recursive, easily implementable procedure for (5-32).

Note that

$$T(\underline{w}_r) = \sum_{j=1}^{n_1/2} (q_1(\alpha_{2j,r}) - q_2(\alpha_{2j-1,r})), \tag{5-33}$$

where

$$q_1(\alpha_{2j,r}) = \sum_{k=(r-1)J+1}^{(r-1)J+\alpha_{2j,r}} K(k) \tag{5-34}$$

$$q_2(\alpha_{2j-1,r}) = \sum_{k=(r-1)J+1}^{(r-1)J+\alpha_{2j-1,r}-1} K(k)$$

and $K(k) \equiv K_o(k) - K_b(k)$. Now let

$$h(\alpha_{2j,r}) = q_1(\alpha_{2j,r}) - 2 \ln p(\alpha_{2j,r}|\hat{\alpha}_{2j-1,r}, \hat{\alpha}_{2j,r-1}) \tag{5-35}$$

and

$$g(\alpha_{2j-1,r}) = -q_2(\alpha_{2j-1,r}) - 2 \ln p(\alpha_{2j-1,r}|\hat{\alpha}_{2j-2,r}, \hat{\alpha}_{2j-1,r-1}), \tag{5-36}$$

where $p(\alpha_{i,r}|\alpha_{i-1,r}, \alpha_{i,r-1})$ is approximated by

$p(\alpha_{i,r}|\hat{\alpha}_{i-1,r}, \hat{\alpha}_{i,r-1})$ and $\hat{\alpha}_{j,r}$ is an estimate of $\alpha_{j,r}$. (5-32) then becomes

$$\begin{aligned} &\min_{\underline{W}} \left\{ \sum_{r=m_{11}}^{m_{21}} \sum_{j=1}^{n_1/2} [h(\alpha_{2j,r}) + g(\alpha_{2j-1,r})] \right\} \\ &= \sum_{r=m_{11}}^{m_{21}} \sum_{j=1}^{n_1/2} [\min_{\alpha_{2j,r}} h(\alpha_{2j,r}) + \min_{\alpha_{2j-1,r}} g(\alpha_{2j-1,r})]. \end{aligned} \tag{5-37}$$

The minimization (5-37) can be performed recursively in a natural way. For $r=m_{11}$, $\hat{\alpha}_{1,m_{11}}$ is obtained by minimizing $g(\alpha_{1,m_{11}})$, which is a function of $\alpha_{1,m_{11}}$ only. Since $h(\alpha_{2,m_{11}})$ obviously involves knowledge of $\hat{\alpha}_{1,m_{11}}$, once $\hat{\alpha}_{1,m_{11}}$ is determined, $\hat{\alpha}_{2,m_{11}}$ can be obtained by minimizing $h(\alpha_{2,m_{11}})$. This process is continuable, allowing us to obtain all the minima involved in (5-37), so that (5-37) is solvable in iterative fashion.

Proceeding any further requires an understanding of the density functions $p(\alpha_{j,r}|\hat{\alpha}_{j-1,r}, \hat{\alpha}_{j,r-1})$. Reasonable density functions are obtainable through our first-order understanding of the component's boundary. For example, suppose that the sloped facet model's output is the quantity $\alpha^{o}_{1,m_{11}}$ for the first boundary point on line m_{11}. Assuming that we want to perturb the boundary using the filter results on a finer grid than that used in the facet model, let us choose a certain multiple x of coarse grid space units as our standard deviation. That is, if Δ is the coarse grid spacing along a row, let $x\Delta$ be the standard deviation, with $\alpha^{o}_{1,m_{11}}$ the assumed mean. We shall then assume a Gaussian density of the form

$$p(\alpha_{1,m_{11}}) = \frac{1}{\sqrt{2\pi}\, x\Delta} \exp\left[- (\alpha_{1,m_{11}} - \alpha^{o}_{1,m_{11}})^2 / 2x^2\Delta^2 \right] \qquad (5\text{-}38)$$

In general, suppose that we have available α^{o}_{ir} for every element i of any row r in the fine grid format. Then choose $x_{ir}\Delta$ units for the standard deviation in such fashion that new choices for $\alpha_{i-1,r}$, α_{ir}, and $\alpha_{i+1,r}$ will follow the appropriate order $\alpha_{i-1,r} \leq \alpha_{ir} \leq \alpha_{i+1,r}$ when the process is completed. Thus a compromise is to be effected between liberty of change and disambiguity of location of boundary points. For example, clearly two boundary points initially $y\Delta$ units apart on a line would require a standard deviation of at

most $y\Delta/6$, so that their adjusted locations would not in all probability cause a reversal of their order on the line. Let us also assume a martingale type process, so that

$$E(\alpha_{ir}|\hat{\alpha}_{i-1,r}, \hat{\alpha}_{i,r-1}) = \hat{\alpha}_{i,r-1}, \tag{5-39}$$

where E denotes the expected value operator. Then we can assume that

$$p(\alpha_{ir}|\hat{\alpha}_{i-1,r}) = \frac{1}{\sqrt{2\pi}\hat{x}_{ir}\Delta} \exp\left[-(\alpha_{ir} - \hat{\alpha}_{i,r-1})^2/2\hat{x}_{ir}^2\Delta^2\right] \tag{5-40}$$

where $\hat{x}_{ir}$ is chosen analogously to x_{ir} in such fashion that a standard deviation of $\hat{x}_{ir}\Delta$ units leads to the proper ordering. That is, we would prefer to use updated values for α_{ir} as they become available rather than the α_{ir}^{o} provided by the first order approximation. Of course, as we examine each section p, where $1 \leq p \leq NS$, we may need to use information provided by the first order methodology, especially when a large number of boundary points per line is involved (corresponding to new branches of the boundary). Note that some of the difficulties inherent in the discussion of the Nahi-Jahanshahi paper are automatically avoided by our a priori knowledge of a reasonable boundary. Once we input (5-38),(5-39), and (5-40) to (5-37), we can solve (5-37) for a given value of m_{11}. A reasonable starting value for m_{11} would be that provided by our first order methodology.

Now that, for a given value of m_{11}, (5-37) has been solved, let us now refer to relation (5-30). Let $\hat{f}(r) = T(\hat{\underline{w}}_r) - 2\ln p(\hat{\underline{w}}_r|\hat{\underline{w}}_{r-1})$, where $\hat{\underline{w}}_r$ has the obvious connotation wherein $\hat{\alpha}_{jr}$ replaces α_{jr}. Since $\hat{f}(r)$ is now a known quantity for every r, consider

$$\min_{m_{11}} \left[-2 \ln p(m_{11}) + \sum_{r=m_{11}}^{m_{21}} \hat{f}(r)\right]. \tag{5-41}$$

The density $p(m_{11})$ can be assumed to be Gaussian with mean m_{11}^{o}, where m_{11}^{o} is the first line of the sloped facet image. Let the standard deviation be $y_1\Delta_1$, where y_1 is some preassigned positive number and Δ_1 is the coarse grid spacing in the column direction. For example, y_1 might be taken to be 1. Then

we have

$$p(m_{11}) = \frac{1}{\sqrt{2\pi}\, y_1 \Delta_1} e^{-(m_{11} - m_{11}^0)^2/2y_1^2\Delta_1^2}. \qquad (5\text{-}42)$$

The minimization problem (5-41) can then be solved through a search procedure. Thus we have a method for analyzing section 1. Sections 2 through NS-1 are resolved in similar fashion, the only difference being that lines m_{1p} and m_{2p} for the pth section, $2 \leq p \leq NS - 1$, are known, so that the only undetermined quantities are the α's. In general, $m_{1p} = m_{2,p-1} + 1$, and $m_{2p} = m_{1,p+1} - 1$. Finally, for section NS, $m_{1,NS}$ is known, but $m_{2,NS}$ is random. Analogously to (5-42), we may use our sloped facet model results to procure a density function for $m_{2,NS}$. If one desires, one may use the results thus obtained to proceed even further. One now can construct a new background process and compare it once more with the object process. Running through the entire operation again, perhaps sweeping the grid in a different order, the boundary can be further refined. After several executions of the method, that boundary corresponding to the largest value of $p(\underline{Y}, \underline{M}, \underline{W})$ could be chosen. A nonlinear filter should produce a measurement model quite similar to (5-9); and, running background and object nonlinear filters, we could use the estimates thus obtained as input to the process just described. We shall not pursue this topic in any further detail.

CHAPTER 6

TARGET IDENTIFICATION AND DESCRIPTION

GENERAL APPROACH

In some sense, the topic of target identification and description is the most interesting of all those items discussed thus far. Much literature has been devoted to this subject, the emphasis being placed on either a geometric or a syntactic understanding of a preprocessed image. For the purposes of target recognition, it seems convenient to adopt a geometric approach of some sort in conjunction with a statistical classification strategy if one desires to understand certain basic target components. On the other hand, syntactic descriptions can be quite useful when one wants to relate different parts of a ship or missile. A hierarchy of understanding is possible in this manner, allowing one, for example, to distinguish enemy from friendly ships. In this chapter, we shall cover both the geometric and syntactic approaches, relying, for our purposes, on methods previously developed in the literature.

Let us begin by emphasizing the importance of multispectral, or multiband, analyses of a target prior to attempting to classify it. We shall illustrate the idea from the point of view of infrared signature processing. Imagine a situation, of sime practical significance, whereby a box arrives in the mail addressed to the President of the United States. It may be that the box contains a nice assortment of shirts and ties, but it is also possible that the box contains a small bomb. By merely looking at the object, one cannot ascertain its contents, but it may well be possible to do so by making use of infrared or sonar spectra. An analogous situation is present in target recognition problems. An engine room is likely to be quite a bit warmer than a passenger space, yet the geometric shape of both compartments could be exactly the same. Therefore, a classifier which explicityly takes into account different wave bands and which attempts to correlate the results in these bands is much more likely to be robust than otherwise.

Let us suppose that we have enumerated a number of generic ship components, each deemed important in some way for recognizing the target. Suppose that there are m of them, designated through the set of labels L_i, $1 \leq i \leq m$. Then the first game which we want to play is that of using the gray level information in order to make an initial assignment of a probability vector $\underline{p}_j = (p_{j1}, p_{j2}, \ldots, p_{jm}, p_{j,m+1})$ to each of n target objects a_j, $1 \leq j \leq n$. In this notation p_{ji}, $1 \leq i \leq m$, is the probability that object a_j has label L_i, $1 \leq i \leq m$. Also, $p_{j,m+1}$ is the probability that a_j has an "unimportant label", which we may denote by L_{m+1}. Now how can one use the gray level information obtained from some component a_j in order to procure the p_{ji}'s? We shall introduce two methods for doing so, the first based on knowledge of object boundary and the second based on gray level information.

BOUNDARY CLASSIFICATION METHODOLOGY

We shall first present a classification strategy based on the boundary, as given by Duda and Hart[12]. It is not a sophisticated approach, since one could very well not tell the difference between two objects having the same boundary, but entirely different temperature profiles. However, it will be seen to have several useful inherent properties. In target recognition one attempts to devise parameters for identification which are, in some sense, independent of orientation in the field of view, position in the field of view, and range from observer to object. Our boundary classifier will automatically have the first two properties, but not the third. In some sense, this is not detrimental when one observes that it is physically impossible anyhow to build a recognition capability totally insensitive to range. All one needs to imagine is two objects of the same shape but of different size which, due to their respective distances from the viewer, cause the same image to be produced. Therefore, what one really tries to do is to make as many parameters insensitive to range as is feasible, since, by so doing, one is bound to reduce computational effort considerably when one wants to recognize an unknown object.

Suppose that we have a discrete description of the boundary and that we have connected the points by a smooth curve (perhaps using cubic spline approximation

methods). Let us look at the curvature k(s) of this approximate boundary, where s is an arc length parameter, and develop it into a Fourier series:

$$k(s) = \sum_{n=-\infty}^{\infty} c_n \exp(2\pi i n s/L), \tag{6-1}$$

where L is the length of the boundary and

$$c_n = \frac{1}{L}\int_0^L k(s)\, \exp(-2\pi i n s/L)ds. \tag{6-2}$$

We shall then agree to use a statistically significant number of the coefficients c_n to build a "feature vector" for the boundary. In other words, one might select a positive integer N and use the vector $(c_{-N}, c_{-N+1}, \ldots, c_0, c_1, \ldots, c_N)$ to characterize the boundary. Note that curvature is an intrinsic property of a smooth curve and thus is <u>automatically</u> independent of orientation or position in a field of view. However, one sees that the curvature of the image boundary varies essentially directly as the distance of the object boundary from the observer. The reason is as follows: The projection of a space curve (X(s), Y(s), Z(s)) onto a viewing plane, as shown in Figure 6-1, is given by[12]

$$\begin{aligned} u(s) &= g(Y(s))\, X(s) \\ v(s) &= g(Y(s))\, Z(s), \end{aligned} \tag{6-3}$$

where

$$g(Y(s)) = f/(Y(s)+f) \tag{6-4}$$

and f is the focal length of the camera employed. Of course, as previously

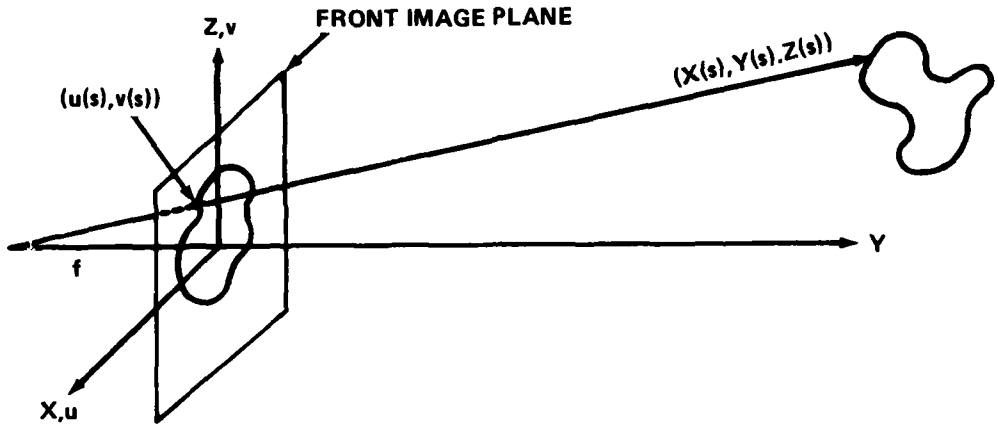


FIGURE 6-1. PROJECTION OF A SPACE CURVE ONTO AN IMAGE CURVE

mentioned, the curve (u(s), v(s)) is only a smooth approximation to the actual boundary curve. It follow that (X(s), Y(s), Z(s)), as related to (u(s), v(s)) through equations (6-3), is really just an approximate representation of the underlying space curve. Now the curvature of the image boundary is representable by the following expression[22]:

$$K_I(s) = \frac{u'(s)v''(s) - v'(s)u''(s)}{\{[u'(s)]^2 + [v'(s)]^2\}^{3/2}}, \qquad (6\text{-}5)$$

where a prime denotes differentiation and where we assume that $u'^2(s)+v'^2(s) \neq 0$ for any s. Forming the derivatives required in (6-5), one finds that

$$K_I(s) = \{2[g'(Y)]^2\,[Y'(s)]^2\,[X(s)\,Z'(s) - X'(s)\,Z(s)]$$

$$+ g(Y)\,g'(Y)\,Y'(s)\,[X(s)\,Z''(s) - X''(s)\,Z(s)]$$

$$+ g(Y)g''(Y)\,[Y'(s)]^2[X'(s)\,Z(s) - X(s)\,Z'(s)]$$

$$+ g(Y)\,g'(Y)Y''(s)[X'(a)\,Z(s) - X(s)\,Z'(s)]$$

$$+ g^2(Y)\,[X'(s)\,Z''(s) - X''(s)\,Z'(s)]\} \qquad (6\text{-}6)$$

$$\{[g'(Y)]^2[Y'(s)]^2[X^2(s) + Z^2(s)]$$

$$+ 2g(Y)g'(Y)Y'(s)\,[X(s)\,X'(s) + Z(s)\,Z'(s)]$$

$$+ g^2(Y)\,[(X'(s))^2 + (Z'(s))^2]\}^{3/2}$$

Note next that $g(Y) \cong f/Y$, $g'(Y) \cong -f/Y^2$, and $g''(Y) \cong 2f/Y^3$. Furthermore, a translation of origin along the Y direction leaves $Y'(s)$ and $Y''(s)$ unchanged. Therefore, factoring $g^2(Y(s))$ from the numerator and $g^3(Y(s))$ from the denominator of (6-6), one has

$$K_I(s) = \frac{O(1/Y)\,F(s) + [X'(s)\,Z''(s) - X''(s)\,Z'(s)]}{g(Y)\,[O(1/Y)G(s) + (X'(s))^2 + (Z'(s))^2]^{3/2}}, \qquad (6\text{-}7)$$

where $F(s)$ and $G(s)$ are well-defined functions of s alone and where the usual order notation has been employed. Now let

$$K_I^A(s) = \frac{1}{g(Y)}\;\frac{X'(s)\,Z''(s) - X''(s)\,Z'(s)}{[(X'(s))^2 + (Z'(s))^2]^{3/2}}, \qquad (6\text{-}8)$$

where we assume that $(X'(s))^2 + (Z'(s))^2 \neq 0$ for any s. If likewise $X'(s)\,Z''(s) - X''(s)\,Z'(s) \neq 0$, one has $K_I^A(s)/K_I(s) \to 1$ as $Y(s) \to +\infty$. Thus, when the camera is translated along the Y axis, $K_I^A(s)$ is asymptotic to $K_I(s)$ at every point where the above requirements are met. Note further that $K_I^A(s)/K_I(s)$ tends <u>uniformly</u> to 1, provided that $X(s)$ and $Z(s)$ are twice continuously differentiable

and that both $(X'(s))^2 + (Z'(s))^2$ and $X'(s)Z''(s) - X''(s)Z'(s)$ are nonzero for every s. Also, we may replace $g(Y(s))$ by $g(Y_{CG})$ for large distances from camera to object, where Y_{CG} is the Y coordinate of the center of geometry. We now have

$$k_i^A(s) \cong \frac{1}{g(Y_{CG})} k_o^A(s), \tag{6-9}$$

where

$$k_o^A(s) \equiv \frac{X'(s)Z''(s) - X''(s)Z'(s)}{[(X'(s)^2 + (Z'(s))^2]^{3/2}}. \tag{6-10}$$

It follows that, for large Y_{CG}, $k_o^A(s)$ can be computed once one has calculated $k_I^A(s)$ and once one knows Y_{CG}. It is assumed that Y_{CG} would be obtained through some tracking algorithm. The expression $k_o^A(s)$ is indeed a range-independent quantity which we can input to (6-1) and (6-2) and for which we can build a feature vector. Once we have such a feature vector, what do we do next? We must perform what is called a target reconstruction. That is, we examine the object from a number of different aspects, and for each aspect i, we construct a feature vector $\underline{c}^i = (c_{-N}^i, c_{-N+1}^i, \ldots, c_o^i, c_1^i, \ldots, c_N^i)$, where $1 \leq i \leq M$ and there are M aspects considered. Assuming that there is a Gaussian process underlying the vectors $\underline{c}^i$, we then form the sample mean vector $\bar{\underline{c}} = \sum_{i=1}^{M} \underline{c}^i/M$ and the sample covariance matrix, whose entries are just $m_{rs} =$

$$\sum_{i=1}^{M} (c_r^i - \bar{c}_r)(c_s^i - \bar{c}_s)/M.$$

THE MOMENT AREA METHOD FOR TARGET CLASSIFICATION

Clearly, the method just presented has built-in limitations. One is forced, first of all, to "regularize" the boundary, i.e., to smooth it in such a manner that corners are eliminated. This may or may not be a proper procedure, depending on the accuracy one wishes to achieve. Secondly, a classifier derived through such boundary considerations may not adequately reflect the underlying gray level signatures. We are thus led to a more sophisticated method called the method of moments, as studied by Hu[2,3] and Dudani[1]. They basically applied the classical theory of moments, as developed by Cayley, to target detection.

Theoretically, an image of an object is uniquely describable in terms of all the moments of its gray levels. From a practical point of view, one needs to form a significant number of such moments. Then one needs to combine them in such manner so as to yield parameters which are invariant with respect to orientation, position in the field of view, and range. Since this approach is well-documented by Hu and Dudani, our discussion in this regard will be cursory.

Before we continue with our discussion of the moment method, let us prodeed a little further with some of the geometrical aspects of target imaging. Such aspects will be important with regard to feature vector formation. The analysis was performed by Dudani in his doctoral dissertation and proceeds as follows: Suppose that, as above, we have a camera coordinate system X, Y, Z as indicated in Figure 6-1. Let us suppose that the axis system x, y, z for our three-dimensional object is initially aligned with this system. Therefore, initially we have

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \tag{6-11}$$

Relation (6-11) provides a fiducial orientation for the object. Having adopted such an orientation, we can specify any new orientation in the following manner: First let us rotate the object about the Y axis through an angle θ. This produces a new coordinate system (x,Y,z) fixed to the object. Next rotate about the new z axis through an angle ψ, producing an (x',y,z) system. Finally rotate about the x' axis through angle ϕ to produce an (x',y',z') coordinate frame. Dudani notes that the product of these three transformations leads to the connection

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \tag{6-12}$$

where we agree, for the sake of notational simplicity, to drop the primes on the object coordinates. Finally one translates the origin of the object frame of

reference to a point (A, B, C) with respect to the camera coordinate system X, Y, Z. One obtains

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \cos\theta\cos\psi & -\cos\theta\sin\psi\cos\phi + \sin\theta\sin\phi & \cos\theta\sin\psi\sin\phi + \sin\theta\cos\phi \\ \sin\psi & \cos\psi\cos\phi & -\cos\psi\sin\phi \\ -\sin\theta\cos\psi & \sin\theta\sin\psi\cos\phi + \cos\theta\sin\phi & -\sin\theta\sin\psi\sin\phi + \cos\theta\cos\phi \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} A \\ B \\ C \end{bmatrix} \quad (6\text{-}13)$$

Suppose next that the optical axis of the camera is directed to pass through the center of geometry (origin) of the object frame, so that A=C=0. Using (6-3) with Y_{CG} replacing Y(s), one sees that

$$u^2 + v^2 = d^2 (X^2 + Z^2), \quad (6\text{-}14)$$

where d is a constant for given Y_{CG} = B. Using (6-13), together with some algebra, one finds that, provided A=C=0, both d and $X^2 + Z^2$ are independent of θ. Therefore, the right side of (6-14) is solely a function of ψ, ϕ, and B. Fixing x, y, z, ψ, ϕ, and B and letting θ vary between 0 and 2π, we find that the point (u,v) traces out a circle in the uv-plane of radius $d(X^2 + Z^2)^{1/2}$. In other words, rotation of the object in the θ variable about its own center of geometry merely produces rotation of the image in the uv-plane. It follows that we need to develop features based on knowledge of ψ and ϕ. On the other hand, if we can obtain features invariant with respect to rotation in uv-coordinates, we may ignore variations in θ.

Let us now go back to our gray scale description in terms of red and blue intensities $g_R(i,j)$ and $g_B(i,j)$ and form the two-dimensional moments of the image

$$m_{pq}^{R} = \frac{1}{g_{TR}} \sum g_R(u_i,v_i)\, u_i^{p}\, v_i^{q}$$

$$m_{pq}^{B} = \frac{1}{g_{TB}} \sum g_B(u_i,v_i)\, u_i^{p} v_i^{q}, \quad (6\text{-}15)$$

where $g_{TR} = \sum g_R(i,j)$, $g_{TB} = \sum g_B(i,j)$, and the sums are over all pixels in the image. The red and blue intensities in (6-15) represent the outputs of a time-varying random process. Typically, they might be the expected values for red and blue signatures under a given set of environmental conditions, which might include solar radiation, cloud cover, atmospheric emissions, and atmospheric haze. In other words they should typify outputs corresponding to actual conditions to which a target would be subjected. Out of (6-15), one can construct, based on the work of Hu and Dudani, the so-called similitude moment invariants with respect to size, elevation, and distance B along the optical axis. Let these be denoted by $M_i^R(\psi, \phi)$ and $M_i^B(\psi,\phi)$, $1 \leq i \leq NI$, where there are NI invariants considered based on the red and blue information. These moment invariants become the elements of our feature vectors. Just as with our boundary vector $(c_{-N}, c_{-N+1},\ldots,c_0, c_1,\ldots,c_N)$, we may consider

$$(M_{1j}^R(\Psi_i,\phi_i), M_{1j}^B(\Psi_i,\phi_i), M_{2j}^R(\Psi_i,\phi_i), M_{2j}^B(\Psi_i,\phi_i),\ldots,M_{NI,J}^R(\Psi_i,\phi_i),$$

$$M_{NI,J}^B(\Psi_i,\phi_i))$$ as a Gaussian vector, where $1 \leq i \leq pa$ covers pa aspects and

$1 \leq j \leq pe$ covers pe sets of environmental conditions. A sample mean and a sample covariance matrix for the pa aspects and the pe environmental situations may then be constructed.

One may imagine selecting a number of interesting objects a_j, $1 \leq j \leq n$, as we mentioned at the beginning of this chapter, with the intent of assigning a label probability vector $\underline{p}_j$ to each, based on our feature vectors (using either a moment area technique or a boundary technique or both). Imagine that we have built a Gaussian distribution for each label L_i, $1 \leq i \leq m+1$. Then, following Dudani,[1] we shall adopt the Bayesian approach, wherein one writes

$$P(C_j|\underline{\rho}) = p(\underline{\rho}|C_j)P(C_j)/p(\underline{\rho}). \qquad (6\text{-}16)$$

Here $p(\underline{\rho}|C_j)$ is the probability density for the feature vector $\underline{\rho}$, given that it belongs to class C_j, $P(C_j)$ is the a priori probability of occurrence of class C_j, and $p(\underline{\rho})$ is the probability density for occurrence of $\underline{\rho}$, regardless of category. Obviously,

$$p(\underline{\varrho}) = \sum_{j=1}^{m+1} p(\underline{\varrho}|C_j)P(C_j), \tag{6-17}$$

where C_j is the class for label L_j. For simplicity, one may assume that $p(C_j) = 1/(m+1)$ for every j if one does not have any better a apriori understanding of the occurrence of classes. In that case (6-16) becomes

$$P(C_j|\underline{\varrho}) = p(\underline{\varrho}|C_j)/\sum_{j=1}^{m+1} p(\underline{\varrho}|C_j). \tag{6-18}$$

In any event the probabilities afforded by (6-16) or (6-18) would become the components of our probability assignment vector $\underline{p}_j$ for object a_j.

RELAXATION METHODOLOGY FOR TARGET CLASSIFICATION

We come now to one of the most important phases of the classification process, namely, that of unifying our knowledge about the individual components of a target in order to understand the target as a whole. This is the purpose of the so-called relaxation methodology which we now introduce. The theory has been developed systematically in papers by Rosenfeld, Hummel, and Zucker, [4] Peleg and Rosenfeld,[23] Zucker, Krishnamurthy, and Haar,[24] Davis and Rosenfeld,[25] and Eklundh, Yamamoto, and Rosenfeld.[5] Now there are two basic types of probabilistic relaxation schemes, which we may call linear and nonlinear, respectively. Both are studied in the paper by Rosenfeld, Hummel, and Zucker just quoted. Therefore, we shall not enter into any elaborate details with regard to convergence except for mentioning that validity of the method has been proved under certain standard conditions for the linear case, but has not been established, in general, for the nonlinear method. Nevertheless experiments have shown that, in practice, the nonlinear procedure works well and yields more realistic results than does the linear one.

We first present the linear relaxation model and show how one adapts it to the target classification problem. The model given by Rosenfeld, Hummel and Zucker is

$$p_i(\lambda) = \sum_j c_{ij}\left[\sum_{\lambda'} p_{ij}(\lambda|\lambda')\, p_j(\lambda')\right], \quad 1 \leq i,j \leq n, \tag{6-19}$$

with the following notational clarification:

(c_{ij}) - doubly stochastic coefficient matrix of positive elements, i.e., one for which each row and column sums to unity.

$p_{ij}(\lambda|\lambda')$ - probability that object a_i has label λ, given that object a_j has label λ'.

$p_i(\lambda)$ - probability that object a_i has label λ.

Also, $p_{ii}(\lambda|\lambda') = 1$ if $\lambda = \lambda'$ and 0 otherwise. Basically one is searching for a fixed point $\{p_i(\lambda)\}$, $1 \leq i \leq n$, for any given label λ. By the Brouwer fixed point theorem [26], such a fixed point always exists. The game is to start with some initial probabilistic labeling $p_j(\lambda)$ for each of our objects a_j and iteratively apply (6-19) until convergence to a stochastic labeling.

Now let us interpret (6–19) in the light of target modeling. First of all, note that the target is viewed at some aspect, say $(\psi_t, \phi_t, \theta_t)$, where the angles are the usual Euler angles of the type we mentioned at the beginning of this chapter. This induces, of course, an Euler representation $(\psi_{tj},\phi_{tj},\theta_{tj})$ for every object a_j. In Dudani's thesis two procedures were introduced for calculating moment invariants, the first of which made use of a principal axis transformation and the second of which produced so-called orthogonal moment invariants. The second procedure would not, in our case, be generally applicable, since it requires that the object be rotated about its own center of geometry. In our case, the entire object, such as a ship, is being rotated about its center of geometry. Clearly, for example, an engine room is not being rotated about its own center of geometry. The analysis given by equations (6-11) - (6-14) still applies, but with the axis of the camera system being directed toward the center of geometry of the ship, missile, etc., involved. One word of caution, however, is in order here. As in any mathematical analysis of a physical situation, one needs to focus his attention on the interrelationships that exist and to be certain that the mathematical theory adequately reflects the real world conditions. If, for example, one is attempting to classify a ship or an airplane, the environment to which it is exposed may have a significant impact on results obtained through infrared, as opposed to visual, processing. There may be a substantial difference in results obtained during the day as opposed to those gathered at night. In the former case, one would normally have to consider

the yaw, roll, and pitch of the target, since each set of angular quantities would correspond to an entirely different radiation pattern. In the latter case, it is expected that environmental factors (such as solar heating) would be minimized, leaving intrinsic sources of energy to govern the signal output. Then it would be possible to show that, if one were, for example, to view the target broadside, pitching motion would not affect the gray level output. That is, any motion in a plane orthogonal to the viewing axis would be unimportant with regard to target component discrimination. Now let us use (6-3) and (6-4) with Y(s) replaced by Y_{CG}, the latter being understood to be the Y coordinate of the entire object's center of geometry. Let us assume that the ship or missile is far enough away so that each point of it can be considered to be at distance Y_{CG} from the camera. Next let ν_{pq} be the pq central moment of the gray levels for the original (X,Z) coordinates and μ_{pq} the corresponding central moments for the image in (u,v) coordinates, which are parallel to the camera coordinates (as shown in Figure 6-1). If there are N pixels in the component of interest, one has

$$\mu_{pq} = \frac{1}{g_{ITR}} \sum_{N \text{ pixels}} g_{IR}(i,j)\ (u_i - \bar{u})^p (v_j - \bar{v})^q, \tag{6-20}$$

where

$$\bar{u} = \sum_{N \text{ pixels}} g_{IR}(i,j) u_i / g_{ITR}$$
$$\bar{v} = \sum_{N \text{ pixels}} g_{IR}(i,j) v_j / g_{ITR}. \tag{6-21}$$

Here $g_{IR}(i,j)$ is the image red signature intensity at pixel (i,j), and $g_{ITR} = \sum_{N \text{ pixels}} g_{IR}(i,j)$. Let (X_{ij}, Z_{ij}) be the X and Z coordinates of an object point which is projected to the image point (u_i, v_j). Also, let

$$\bar{X} = \sum_{n \text{ pixels}} g_{tR}(i,j) X_{ij} / g_{tTR}$$
$$\bar{Z} = \sum_{n \text{ pixels}} g_{tR}(i.j) Z_{ij} / g_{tTR}, \tag{6-22}$$

where $g_{tR}(i,j)$ is the target red signature intensity for surface pixel (i,j) and $g_{tTR} = \sum_{N\ pixels} g_{tR}(i,j)$. Now $u_i = dX_{ij}$, $\bar{u} = d\bar{X}$, $v_j = dZ_{ij}$, and $\bar{v} = d\bar{Z}$, where $d = f/(Y_{CG} + f)$. Furthermore, assume that there exists a red attenuation factor $a_R(d)$, solely a function of the distance traversed from object to camera (or detector) by the red component output. Then, for every (i,j),

$$g_{IR}(i,j) = a_R(d)g_{tR}(i,j). \tag{6-23}$$

Substituting (6-23) into (6-20), we find that

$$\mu_{pq} = d^{p+q}\nu_{pq}, \tag{6-24}$$

where

$$\nu_{pq} = \frac{1}{g_{tTR}} \sum_{N\ pixels} g_{tR}(i,j)(X_{ij} - \bar{X})^p(Z_{ij} - \bar{Z})^q. \tag{6-25}$$

Note that (6-24) does not involve the attenuation factor $a_R(d)$. Also, we are assuming, in our analysis, that radiation propagates along straight-line paths, i.e., that any refraction phenomena are of secondary significance. Now, using the fact, from (6-24), that $\mu_{20} = d^2\nu_{20}$ and eliminating d^2, one sees that $\mu_{pq}/\mu_{20}^{(p+q)/2}$ is a size invariant. That is to say, if $\mu_{pq}^{(1)}$ is the result of using d_1 in (6-24) and $\mu_{pq}^{(2)}$ is the result of using d_2, one has

$\mu_{pq}^{(1)}/(\mu_{20}^{(1)})^{(p+q)/2} = \mu_{pq}^{(2)}/(\mu_{20}^{(2)})^{(p+q)/2}$. Furthermore, we would like to use for μ_{pq} certain fiducial quantities, namely, central moments with respect to principal axes. Assuming that our origin in the image plane is at the red intensity center of geometry of the component image and employing the rotation of coordinates

$$\begin{aligned} u' &= u\cos\theta - v\sin\theta \\ v' &= u\sin\theta + v\cos\theta, \end{aligned} \tag{6-26}$$

as Dudani does, one finds that

$$2\mu_{11}' = (\mu_{20} - \mu_{02})\sin 2\theta + 2\mu_{11}\cos 2\theta. \tag{6-27}$$

By definition, the principal axes are those for which $\mu_{11}' = 0$. This leads to the condition

$$\tan 2\theta = -2\mu_{11}/(\mu_{20} - \mu_{02}). \qquad (6\text{-}28)$$

The angle θ, as given by (6-28), is the orientation of the principal axis system relative to our uv coordinate frame (parallel to the camera frame). Our invariants to size and rotation then become $M_{pq}' = \mu_{pq}'/(\mu_{20}')^{(p+q)/2}$, provided that $(p,q) \neq (2,0)$. Note that M_{pq}' may be computed directly without any knowledge of distance from object to image. There is one more invariant, namely, μ_{20}'/d^2, which requires a knowledge of d for its determination. Through this invariant one gains an understanding of target component size. Of course one can perform the same sequence of computations for the blue intensities.

When the objects a_j are projected into the image plane of a camera, they will, of course, adopt certain positions in that plane relative to one another. The probabilities $p_{ij}(\lambda|\lambda')$ in (6-19) represent compatabilities that must exist among such objects and, as such, should be prespecified. Suppose now that one is addressing k target types, say $\{T_i\}$, $1 \leq i \leq k$. Suppose, also, that, for each target type T_i, we build a training set for a number of target aspects $(\theta_t^{ri}, \psi_t^{ri}, \phi_t^{ri})$, $1 \leq r \leq p$, there being p overall views considered. Such training sets, inasmuch as they might be employed to derive distance relationships among components, could be obtained under idealized conditions. For example, the target itself, say a ship, could be studied as a model in a laboratory setting if one's only interest is to procure the distance between the engine room and passenger quarters at any given target aspect. A histogram analysis could then be performed over all the ship types T_i, $1 \leq i \leq k$, and numbers $p_{ij}(\lambda|\lambda';D)$ could be derived over all ship types and aspects. Here $p_{ij}(\lambda|\lambda';D)$ means specifically the probability that object a_i has label λ, given that object a_j has label λ' and is at distance D from a_j. There is an alternative point of view which could be taken at this point which should be mentioned. Note that, on calculating $p(\rho|C_j)$ for input to (6-16), we have identified C_j with the jth ship type, so that C_j is really T_j. In effect, we have regarded all out target image feature vectors, regardless of target orientation, as coming from the same Gaussian distribution. In parallel with this concept, it is reasonable to calculate $p_{ij}(\lambda|\lambda';D)$ based on all aspects of

the target. A better way of approaching the task perhaps is that taken by Dudani. Instead of identifying C_j solely with T_j, one includes in C_j the target aspect (θ_t,ψ_t,ϕ_t), so that the basic Gaussian density $p(\underline{\rho}|C_j)$ has a different statistical interpretation. Feature vectors are constructed under varying environmental conditions, but for a given target aspect and target type. In the spirit of Dudani's approach, the prior probability $P(C_j)$ in (6-16) is

$$P(C_j) = P(T_j)p(\psi_t|\psi_{t_o})p(\phi_t|\phi_{t_o})p(\theta_t|\theta_{t_o}), \tag{6-29}$$

wherein $P(T_j)$ is the probability of the target type T_j, $p(\psi_t|\psi_{t_o})$ is the Gaussian density for yaw ψ_t, given a nominal value ψ_{t_o}, $p(\phi_t|\phi_{t_o})$ is the Gaussian density for roll ϕ_t, given nominal ϕ_{t_o}, and, finally, $p(\theta_t|\theta_{t_o})$ is the density for pitch θ_t, given nominal pitch θ_{t_o}. Likewise we should now understand $p_{ij}(\lambda|\lambda';D)$ a little differently. We should again perform a histogram analysis, but condition it on target aspect. One then needs to integrate (6-19) over all target aspects, i.e., one considers

$$\int_{-\pi/2}^{\pi/2}\int_{0}^{2\pi}\int_{0}^{2\pi} p_i(\lambda|\theta_t,\psi_t,\phi_t)p(\theta_t|\theta_{t_o})p(\psi_t|\psi_{t_o})p(\phi_t|\phi_{t_o})d\theta_t d\psi_t d\phi_t \tag{6-30}$$

$$= \sum_j c_{ij}\Big[\sum_{\lambda'}\int_{-\pi/2}^{\pi/2}\int_{0}^{2\pi}\int_{0}^{2\pi} p_{ij}(\lambda|\lambda';D,\theta_t,\psi_t,\phi_t)p_j(\lambda'|\theta_t,\psi_t,\phi_t)p(\theta_t|\theta_{t_o})$$
$$\cdot p(\psi_t|\psi_{t_o})p(\phi_t|\phi_{t_o})d\theta_t d\psi_t d\phi_t\Big]$$

If the Gaussian (or, possibly, uniform) random variables θ_t, ψ_t, and ϕ_t in (6-29) have small variances, then we may approximate $p_{ij}(\lambda|\lambda'; D,\theta_t,\psi_t,\phi_t)$ by $p_{ij}(\lambda|\lambda';D,\theta_{t_o},\psi_{t_o},\phi_{t_o})$. In that case (6-30) becomes

$$p_i(\lambda|\theta_{t_o},\psi_{t_o},\phi_{t_o}) = \sum_j c_{ij}\left[\sum_{\lambda'} p_{ij}(\lambda|\lambda';D,\theta_{t_o},\psi_{t_o},\phi_{t_o})\right. \tag{6-31}$$

$$\left.\cdot p_j(\lambda'|\theta_{t_o},\psi_{t_o},\phi_{t_o})\right]$$

This mechanism is theoretically perhaps more rigorous, but, at the same time, may require much more work effort to implement.

Either (6-19) or (6-31) could, in principle, be used in an iterative scheme for procuring consistent labeling vectors $(p_j(\lambda))$ for the components a_j. Having thus identified the components, together with their relative positions, one may, at a higher level, be able to say, with some degree of certainty, what type of target is present. If the target changes heading or if we permit ourselves to walk around the target in order to improve our understanding of it and rerun the basic iterative scheme, there is a good chance that we can improve our understanding of the category of target involved.

We would also like to say that certain topological properties might be useful in target classification.[27] For example, is the object a_j we are considering convex? Is it connected? If it is connected, what can be said about the number of holes contained in the image? Such higher level information could certainly be used as a logical input to classification of components of a target. This prior knowledge could indeed be utilized in defining $p(T_j)$ as input to (6-29).

Let us now proceed with the nonlinear probabilistic model suggested by Rosenfeld, Hummel and Zucker. It turns out that such a model is more useful than the linear model in two different ways: (1) It appears to converge, in a large number of practical applications, to a limit not depending on initial assignment probabilities. (2) It exhibits favorable compatibility properties with regard to generation of $p_i(\lambda)$ consistent with correlations between objects. The model is predicated on certain known correlations rather than compatibilities. The correlations are denoted by $r_{ij}(\lambda,\lambda')$, where, of course, $|r_j(\lambda,\lambda')| < 1$. One defines a quantity $q_i^{(k)}(\lambda)$ analogous to the $p_i^{(k)}(\lambda)$ (the kth iterate of $p_i(\lambda)$) defined in the linear model:

$$q_i^{(k)}(\lambda) = \sum_j c_{ij}\Big[\sum_{\lambda'} r_{ij}(\lambda,\lambda')p_j^{(k)}(\lambda')\Big], \tag{6-32}$$

where again the c_{ij} are nonnegative and the matrix (c_{ij}) is doubly stochastic. Then we define

$$p_i^{(k+1)}(\lambda) \equiv p_i^{(k)}(\lambda)[1+q_i^{(k)}(\lambda)] \Big/ \sum_{\lambda} p_i^{(k)}(\lambda)[1+q_i^{(k)}(\lambda)]. \tag{6-33}$$

It turns out that $|q_i^{(k)}(\lambda)| < 1$, so that, from (6-33), $0 < p_i^{(k+1)}(\lambda) < 1$. Again, we can certainly invoke a model similar to our linear version in the case where aspect distinctions are not made. However, when they are made, as in the case (6-26) just studied, it now becomes impossible to parallel the modeling in (6-30), which goes through because of the linearity of the operator. On the other hand if we assume that we <u>know</u> the values of ψ and ϕ, i.e., if $\psi = \psi_0$, $\phi = \phi_0$ and $\theta = \theta_0$, then we need not average as we do in (6-30), and one merely applies (6-33), given the aspect (θ_0,ψ_0,ϕ_0). The problem we now pose is that of providing the inputs $r_{ij}(\lambda,\lambda')$. For a given aspect, we suppose that $r_{ij}(\lambda,\lambda')$ depends on the distance D between the image of object a_i and the image of object a_j. Following Rosenfeld, Hummel and Zucker, let A be the event that object a_i has label λ, and let B be the event that a_j has label λ'. Then

$$r_{ij}(\lambda,\lambda';D) = \frac{p\ (A \text{ and } B;D) - p(A)p(B)}{[(p(A) - p^2(A))\ (p(B) - p^2(B))]^{1/2}}.$$

The quantity p(A and B;D) is the probability that a_i has label λ and a_j has label λ' when it is observed that they are separated by D units in the image plane. We can compute, as before, a frequency of joint occurrence of labels λ and λ' for objects a_i and a_j separated by D units, where $0 < D < \infty$. Also , we can calculate the frequency of occurrence of a given label λ over all target types and input that quantity for p(A). A similar statement pertains for computing p(B).

OPTIMIZATION OF STOCHASTIC LABELING

A scheme presented by Peleg[28] is of great importance in deciding upon the merit of a stochastic labeling and, subsequently, making a specific labeling assignment. Peleg assumes that one wants to maximize the product of two probabilities, one of which he calls the stochastic probability $P_S(\Omega)$ for a given labeling Ω of nodes of a network and the other of which $P_M(\Omega)$ he calls the model probability corresponding to the labeling Ω. The function to be optimized is

$$P(\Omega) = P_S(\Omega)P_M(\Omega). \qquad (6\text{-}34)$$

Suppose that, after executing a relaxation procedure as previously described, we obtain a stochastic labeling for every object a_j, i.e., we have a label probability assignment vector for each a_j. Then choose that label at a_j for which the probability is maximum, and let $\Omega = (\alpha_1,\alpha_2,\ldots,\alpha_n)$, where n is the number of objects and α_i is the label for object a_i yielding maximum probability. Let $P_S(\Omega)$ be the product of these n maximum probabilities. Next form the quantity $P_M(\Omega)$, which is the a priori probability for the given label assignment. This can be obtained, at least approximately, using the compatibilities previously developed (by histogram techniques). Now one can adopt the following strategy: One may choose not to iterate (6-31) or (6-33) to a fixed point, but instead monitor the output, stopping at that point where (6-34) is maximum. Also, one may employ several viewer aspects in conjunction with (6-31), (6-33), and (6-34), again choosing that Ω for which (6-34) is maximum. In this way one procures a specific labeling from which one can better identify the target.

CHAPTER 7

SUMMARY

The intent of this report has been to attempt a unification of ideas on image modeling presented in recent years. Also, it is felt that originality has been injected into the basic discrimination procedures, so that more meaningful directions may be pursued. We have emphasized the early processing of information, since that is of paramount importance. The labeling evaluation procedures given in Chapter 6 would certainly be meaningless unless one can construct reasonable feature vectors as prototypes of the classes. One of the main purposes of the work would be to make available a more sophisticated methodology for target recognition, so that one could ultimately distinguish enemy from friendly targets.

In the future it is hoped that this work can be interfaced in some meaningful way with data from realistic scenarios. Hopefully a computer code can be developed around the ideas presented here.

REFERENCES

1. Dudani, Sahibsingh Amulsingh, An Experimental Study of Moment Methods for Automatic Identification of Three-Dimensional Objects from Television Images, Doctoral Dissertation, Ohio State University, 1973.

2. Hu, M. K., Identification of Three-Dimensional Aircraft by Their Two-Dimensional Optical Images, Syracuse University, RADC-TR-75-201, Vol. I, Aug 1975.

3. Hu, M. K., Theory of Moment Invariants and Applications to Adaptive Pattern Recognition; Theory of Adaptive Mechanisms, Syracuse University, RADC-TDR-63-334, Dec 1963.

4. Rosenfeld, Azriel, Hummel, Robert A., and Zucker, Steven W., "Scene Labeling by Relaxation Operations," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-6, No. 6, Jun 1976, pp. 420-433.

5. Eklundh, J. O., Yamamoto, H., and Rosenfeld, A., "A Relaxation Method for Multispectral Pixel Classification," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-2, No. 1, Jan 1980, pp. 72-75.

6. Haralick, Robert M., "Edge and Region Analysis for Digital Image Data," in Image Modeling, Azriel Rosenfeld, ed., Academic Press, 1981, pp. 171-184.

7. Nahi, Nasser E. and Jahanshahi, Mohammed H., "Image Boundary Estimation," IEEE Transactions on Computers, Vol. C-26, No. 8, Aug 1977, pp. 772-781.

8. Kaufman, Howard, Woods, John W., Dravida, Subrahmanyam, and Tekalp, A. Murat, "Estimation and Identification of Two-Dimensional Images," IEEE Transactions on Automatic Control, Vol. AC-28, No. 7, Jul 1983, pp. 745-756.

9. Woods, John W., "Kalman Filtering in Two Dimensions, "IEEE Transactions on Information Theory, Vol. IT-23, No. 4, Jul 1977, pp. 473-482.

10. Chen, P. C. and Pavlidis, T., "Image Segmentation as an Estimation Problem," in Image Modeling, Ariel Rosenfeld, ed., Academic Press, 1981, pp. 9-28.

11. Samet, Hanan, Connected Component Labeling Using Quadtrees, University of Maryland, TR-756, Apr 1979.

12. Duda, Richard O. and Hart, Peter E., Pattern Classification and Scene Analysis, (New York: Wiley, 1973).

13. Fisz, Marek, Probability Theory and Mathematical Statistics, Third Edition, (New York: Wiley, 1963).

14. Papoulis, Athanasios, Probability, Random Variables and Stochastic Processes, (New York: McGraw-Hill, 1965).

15. Cohen, Edgar A., Jr., Simultaneous Approximation of Position and Velocity by Regressions Linear in Their Parameters, NOLTR 74-37, 1 Mar 1974.

16. Cohen, Edgar A., Jr., CCAMIN, A Computer Program for Fitting Time of Arrival and Shock Peak Overpressure Data in a Gas, NOLTR 72-169, 4 Aug 1972.

17. Hogg, Robert V. and Craig, Allen T., Introduction to Mathematical Statistics, 3rd Edition, (London: MacMillan, 1970).

18. Davis, Philip J., Interpolation and Approximation, (New York: Blaisdell 1963).

19. Feller, William, An Introduction to Probability Theory and Its Applications, Vol. II, (New York: Wiley, 1966).

20. Saridis, G. N., "Comparison of Six Online Identification Algorithms," Automation, Jan 1974, pp. 69-79.

21. Nahi, N. E., and Habibi, A., "Nonlinear Adaptive Recursive Image Enhancement," in Proc. 1st Int. Joint Conf. Pattern Recognition, Oct 1973, pp. 485-491.

22. Randolph, John F. and Kac, Mark, Analytic Geometry and Calculus, (New York: MacMillan, 1954).

23. Peleg, Shmuel and Rosenfeld, Azriel, "Determining Compatibility Coefficients for Curve Enhancement Relaxation Processes," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-8, No. 7, Jul 1978, pp. 548-555.

24. Zucker, Steven W., Krishnamurthy, E. V., and Haar, Robert L., "Relaxation Processes for Scene Labeling: Convergence, Speed, and Stability," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-8, No. 1, Jan 1978, pp. 41-48.

25. Davis, L. S., and Rosenfeld, A., "Applications of Relaxation Labeling, 2:Spring-Loaded Template Matching," Proc. 3rd Int. Joint Conf. on Pattern Recognition, Nov 1976, pp. 591-597.

26. Karlin, Samuel, Mathematical Methods and Theory in Games, Programming, and Economics, (Reading,MA: Addison-Wesley, 1959).

27. Tou, J. T., "Feature Extraction in Pattern Recognition, "Pattern Recognition, Pergamon Press, Vol. 1, 1968, pp. 3-11.

28. Peleg, Shmuel, Monitoring Relaxation Algorithms Using Labeling Evaluation, TR-842, University of Maryland, Dec 1979.

DISTRIBUTION

	Copies
RCA-Astroelectronics Attn: Mr. Kent Anderson T2G P. O. Box 800 Princeton, NJ 08540	1
Dr. Dan Love 666 S.W. Carrera Lane Lake Oswego, OR 97034	1
Mitre Corporation Attn: Mr. E. Gordon Powell 1820 Dolley Madison Drive McLean, VA 22102	1
University of Maryland Attn: Prof. A. Rosenfeld Computer Science Center College Park, MD 20742	1
Princeton University Attn: Prof. P. C. Chen Dept. of Electrical Engineering and Computer Science Princeton, NJ 08540	1
University of Maryland Attn: Prof L. S. Davis Computer Sciences Dept. College Park, MD 20742	1
Purdue University Attn: Prof. K. S. Fu School of Electrical Engineering W. Fayette, IN 47907	1
George Washington University Attn: Prof. M. Lowe Computer Sciences Dept. 801 22nd St., N.W. Washington, D.C. 20052	1
Virginia Polytechnic Institute and State University Attn: Prof. R. M. Haralick Department of Electrical Engineering and Department of Computer Science Blacksburg, VA 24061	1
University of Arizona Attn: Prof. B. R. Hunt Systems Engineering Dept. and Optical Sciences Center Tuscon, AZ 85721	1
SRI International Attn: Prof. J. M. Tenenbaum Artificial Intelligence Center Menlo Park, CA 94025	1
University of Florida Attn: Prof. Julius T. Tou Center for Information Research Gainesville, FL 32611	1
McGill University Attn: Prof. Steven W. Zucker Computer Vision & Graphics Lab. Department of Electrical Engineering Montreal, Quebec, Canada	1
University of Maryland-Balti. Co. Attn: Dr. A. Brooke Stephens Math & Psychology Building 4th Floor 5401 Wilkins Ave. Catonsville, MD 21228	1
Department of the Navy Naval Air Systems Command Attn: Dr. G. F. Heiche, AIR 03D 424 JP1 Washington, D. C. 20361	1
Office of Naval Research Attn: Dr. E. J. Wegman 607 BCT1 Arlington, VA 22217-500	1

	Copies
Defense Technical Information Center Cameron Station Alexandria, VA 22314	12
Library of Congress Attn: Gift & Exchange Div. Washington, D. C. 20540	4
Internal Distribution	
E06 (Moscar, Caviris)	2
N15 (Rowan)	1
N21 (Masters)	1
N51 (Price)	1
R14 (McDonald)	1
R40 (Blatstein)	1
R42 (Caswell, Wilson)	2
R42 (Ostrowski, Kessler)	2
R42 (Hall, Kaufman, Katz)	3
R43 (Scarzello)	1
R44 (Van Tuyl, Wingate)	2
R44 (Cohen)	12
U22 (Ballard)	1
E231	9
E232	3

END

FILMED

10-85

DTIC